

### (19) World Intellectual Property Organization International Bureau





## (43) International Publication Date 12 April 2001 (12.04.2001)

#### **PCT**

## (10) International Publication Number WO 01/24781 A2

(51) International Patent Classification7: A61K 31/00

(21) International Application Number: PCT/CA00/01188

(22) International Filing Date: 6 October 2000 (06.10.2000)

(25) Filing Language:

English

(26) Publication Language:

English

(30) Priority Data:

 2,285,690
 7 October 1999 (07.10.1999)
 CA

 60/158,043
 7 October 1999 (07.10.1999)
 US

 60/217,765
 12 July 2000 (12.07.2000)
 US

- (71) Applicant (for all designated States except US): NOVA-NEURON INC. [CA/CA]; 5859 University Avenue, Sir Charles Tupper Medical Bldg., Room 15D7, Halifax, Nova Scotia B3H 4H7 (CA).
- (72) Inventors; and
- (75) Inventors/Applicants (for US only): ROBERTSON, Harold, A. [CA/CA]; 2384 Clifton Street, Halifax, Nova Scotia B3K 4V1 (CA). DENOVAN-WRIGHT, Eileen, M. [CA/CA]; 22 Fleming Drive, Halifax, Nova Scotia B3P 1A9 (CA).

- (74) Agent: HELLER, David, J.; Ridout & Maybee, One Queen Street East, Suite 2400, Toronto, Ontario M5C 3B1 (CA).
- (81) Designated States (national): AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CR, CU, CZ, DE, DK, DM, DZ, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZW.
- (84) Designated States (regional): ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

#### Published:

 Without international search report and to be republished upon receipt of that report.

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.



A2

(54) Title: GENE NECESSARY FOR STRIATAL FUNCTION, USES THEREOF, AND COMPOUNDS FOR MODULATING SAME

(57) Abstract: PDE10A, a gene that is normally highly expressed in mammalian striatum and elsewhere, has been found to decrease in expression during the development of CAG repeat disorders such as Huntington's disease. The invention teaches a method for detecting the presence of or the predisposition for a CAG repeat disorder. Compounds which modulate CAG repeat disorders and their uses are taught. Methods for screening for further compounds to modulate CAG repeat disorders are also taught.

# Gene Necessary for Striatal Function, Uses Thereof, and Compounds for Modulating Same

#### **CROSS-REFERENCE**

This patent claims priority from Canadian Patent application no. 2,285,690 filed October 7, 1999, US provisional application no. 60/158,043 filed October 7, 1999, and US provisional application no. 60/217,765 filed July 12, 2000, entitled Gene Necessary for Striatal Function, Uses Thereof, and Compounds for Modulating Same.

#### FIELD OF THE INVENTION

The present invention relates to a polynucleotide, PDE10A, which is down-regulated during the development of CAG repeat disorders, such as Huntington's disease. The present invention also describes compounds that modulate CAG repeat disorders, processes for expressing PDE10A, and its agonists and antagonists, and uses of PDE10A, and its variants, derivatives, agonists and antagonists.

#### BACKGROUND OF THE INVENTION

Very few if any effective treatments exist for neurological disorders characterized by progressive cell loss, known as neurodegenerative diseases, as well as those involving acute cell loss, such as stroke and trauma.

Huntington's disease (HD) is an inherited neurological disorder that is transmitted in

1

autosomal dominant fashion. HD results from genetically programmed degeneration of neurons in certain areas of the brain. Huntington's disease is caused by a mutation of the gene IT-15 that codes for the protein huntingtin. The huntingtin gene contains a polymorphic stretch of repeated CAG trinucleotides that encode a polyglutamine tract within huntingtin. If this tract exceeds 35 in number, Huntington's disease results. Huntington's disease is only one of a number of neurological diseases which are characterised by these polyglutamine repeats (Ross, 1997). Schizophrenia, Alzheimer's disease, stroke, trauma, and Parkinson's disease also affect the basal ganglia.

Huntingtin has no sequence similarity to known proteins (Group THDCR, 1993; Sisodia, 1998). The function of the normal or mutated HD form of huntingtin has not been defined by the prior art. It is evident, however, that the expression of the HD form of huntingtin leads to progressive and selective neuronal loss. It has been demonstrated that the GABA- and enkephalin-containing medium spiny projection neurons of the caudate-putamen eventually die as a result of HD (Richfield et al., 1994). Patients with minimal cell loss, however, still present with motor and cognitive symptoms suggesting that neuronal dysfunction, and not simply cell loss, contribute to the symptoms of HD. The motor symptoms of HD include the development of chorea, dystonia, bradykinesia and tremors (Young et al., 1986). Voluntary movements may also be affected such that there may be disturbances in speech (Ludlow et al., 1987) and degradation of fine motor co-ordination (Young et al., 1986). In addition to motor decline, emotional disturbances and cognitive loss are also evident during the progression of HD (Caine et al., 1978).

Despite the fact that huntingtin is ubiquitously expressed, HD specifically affects cells of the

basal ganglia, structures deep within the brain that have a number of important functions, including co-ordinating movement. The basal ganglia includes the caudate nucleus, the putamen, the nucleus accumbens and the olfactory tubercule. HD also affects the brain's outer surface, or cortex, which controls thought, perception, and memory. The mechanism by which only a small group of neurons in the striatum and cortex are rendered vulnerable to this ubiquitously expressed mutant protein is not known. There are no effective treatments for Huntington's disease.

Huntington's disease is widely believed to be a gain-of function disorder but neither the normal function nor the gained function of huntingtin is known. Because the function for huntingtin is not known, there is little insight into the disease process. It was believed that huntingtin was related to neuronal intranuclear inclusions (NII). However, recent results have cast doubt on our understanding of the role of the NII in Huntington's disease (Saudou et al., 1998) or in other CAG repeat disorders (Klement et al., 1998; see also commentary by Sisodia, 1998).

The development of a mouse carrying the 5' end of the human Huntington's disease gene (the promoter and first exon; Mangiarini et al., 1996) was an important step in the development of the tools that will allow us to understand the function (and gain-of-function) associated with huntingtin. R6/2 mice exhibit a rapidly progressing neurological phenotype with onset at about 8 weeks. This phenotype includes a movement disorder characterised by shuddering, resting tremor, epileptic seizures and stereotyped behaviour. These symptoms suggest that the function of the basal ganglia is affected by the expression of the human exon 1 transgene prior to neuronal cell death. By 12 weeks the affected mice have significantly reduced brain

weights and they die by about 13 weeks of age. Neuronal intranuclear inclusions (NII) develop at about 4 weeks (Davies et al., 1997). As is observed in human Huntington's disease patient, the R6/2 mice show changes in neuronal receptors (Cha et al., 1998). The present inventors have also demonstrated that changes in the expression of DARPP-32 and cannabinoid receptors change over time in HD mice; such changes have also been observed in human Huntington's disease patients (unpublished results). The loss of the cannabinoid receptor is one of the earliest documented changes that occur prior to neuronal degeneration in human HD patients. The R6/2 model, therefore, mimics the early phases of HD; a point in disease development where intervention would be most appropriate.

Human PDE10 was recently identified by identification of cDNA fragments published on the National Center for Biotechnology Information (NCBI) Expressed Sequence Tags (EST) database (Loughney et al., WO99/42596). While PDE10 was found to share homology with known PDEs, no function could be identified for PDE10.

#### SUMMARY OF THE INVENTION

The present invention provides the function and uses of a nucleotide segment, PDE10A, and compounds which inhibit or promote the development of CAG repeat disorders such as Huntington's Disease.

The invention teaches a method for identifying a compound which inhibits or promotes a CAG repeat disorder, comprising the steps of: (a) selecting a control animal having PDE10A and a test animal having PDE10A; (b) treating said test animal using a compound; and (c)

Ļ

determining the relative quantity of RNA corresponding to PDE10A, as between said animals. In an embodiment, the animal is a mammal, preferably a mouse, and preferably a transgenic mouse. In an embodiment, the CAG repeat disorder is Huntington's disease.

The invention also teaches a method for identifying a compound which inhibits or promotes a CAG repeat disorder, comprising the steps of: (a) selecting a host cell containing PDE10A; (b) cloning said host cell and separating said clones into a test group and a control group; (c) treating said test group using a compound; and (c) determining the relative quantity of RNA corresponding to PDE10A, as between said test group and said control group. In an embodiment, the CAG repeat disorder is Huntington's disease.

The invention further teaches a method for detecting the presence of or the predisposition for a CAG repeat disorder, said method comprising determining the level of expression of RNA corresponding to PDE10A in an individual relative to a predetermined control level of expression, wherein a decreased expression of said RNA as compared to said control is indicative of a CAG repeat disorder. Preferably, the expression is measured by in situ hybridization, fluorescent in situ hybridization, polymerase chain reaction, or DNA fingerprinting technique. In an embodiment, the CAG repeat disorder is Huntington's disease.

The invention further teaches compositions for treating a CAG repeat disorder comprising a compound which modulates PDE10 expression and a pharmaceutically acceptable carrier.

The compound can be selected from the group consisting of: quinpirole, alloxan, miconazole nitrate, MDL-12330A and tetracyline derivatives such as demeclocycline. The compound

may be selected from the group consisting of: (6R,12aR)-2,3,6,7,12,12a-Hexahydro-6-(5-benzofuranyl)-2-methyl-pyrazino[2', 1':6,1]pyrido[3,4-b]indole-1,4-dione,

(6R,12aR)-2,3,6,7,12,12a-Hexahydro-6-(5-benzofuranyl)-pyrazino[2',1':6,1]pyrido[3,4]indole-1,4-dione, (6R,12aR)-2,3,6,7,12,12a-Hexahydro-6-(5-benzofuranyl)-2-isopropylpyrazino[2',1':6,1]pyrido[3,4-b]indole-1,4-dione, (3S,6R,12aR)-2,3,6,7,12,12a-Hexahydro-6(5-benzofuranyl)-3-methyl-pyrazino[2',1':6,1]pyrido[3,4-b]indole-1,4-dione, and

(3S,6R,12aR)-2,3,6,7,12,12a-Hexahydro-6-(5-benzofuranyl)-2,3-dimethyl-pyraz

ino[2',1':6,1]pyrido[3,4-b]indole-1,4-dione, or from the group consisting of: KS-505,

IC224,SCH 51866, IBMX and Dipyridamole. The disorder can be HD.

The invention also teaches the use of a composition which modulates PDE10 for treating a CAG repeat disorder comprising administering the composition to a subject in need of such treatment, and such use of the composition which modulates PDE10 for treating HD.

#### BRIEF DESCRIPTION OF THE DRAWINGS

. 4

FIG. 1 is a portion of an autoradiogram of the differential display reaction identifying PDE10A in mouse brain mRNA.

FIG. 2 is a northern blot confirming that PDE10A has a lower steady-state level of expression in the striatum of transgenic HD mice.

FIG. 3 is a nucleotide sequence of the differential display cDNA fragment of pPDE10A.

FIG. 4 shows the *in situ* hybridization of probe 1 to coronal and saggital brain sections of 10 week-old wild-type and HD mice.

FIG. 5 shows the *in situ* hybridization corresponding to spatial and temporal expression of PDE10A in brain sections of wild-type and HD mice over the period of time that the HD mice develop abnormal movements and postures.

FIG. 6 shows the *in situ* hybridization corresponding to expression of PDE10A in brain sections of one day old wild-type and HD mice.

FIG. 7 shows the *in situ* hybridization corresponding to distribution of the mRNA of PDE10A in mouse striatal neurons.

FIG. 8 is the *in situ* hybridization corresponding to mRNA distribution of the rat homologue of PDE10A in rat brain tissue.

FIG. 9 shows a Southern blot analysis of DNA from wild-type and transgenic HD mice hybridized to the pPDE10A cDNA probe.

FIG. 10 is a nucleotide sequence of cPDE10-1, and corresponds to SEQ ID NO. 1.

FIG. 11 is a restriction map of cPDE10-1.

Ļ

FIG. 12 is a nucleotide sequence of cPDE10-2, and corresponds to SEQ ID NO. 2.

FIG. 13 is a restriction map of cPDE10-2.

FIG. 14 is a schematic diagram showing the alignment of cPDE10-1 and -2 and the regions that are identical and unique between the two clones.

FIG. 15 is a nucleotide sequence of cPDE10A and RACEs, corresponding to SEQ ID NO. 11.

FIG. 16 is a map of PDE10A coding sequence and restriction sites.

FIG. 17 is a map of PDE10A coding sequence and features.

FIG. 18 is a restriction map of PDE10A.

FIG. 19 is a nucleotide sequence of cPDE10A and corresponds to SEQ ID NO. 12.

DETAILED DESCRIPTION OF EMBODIMENTS OF THE INVENTION

The following illustrative explanations are provided to facilitate understanding of certain terms used frequently herein. The explanations are provided as a convenience and are not limitative of the invention.

"Host cell" is a cell which has been transformed or transfected, or is capable of transformation or transfection by an exogenous polynucleotide sequence.

"Identity", "similarity" or "homologous", as used in the art, are relationships between two or

more polynucleotide sequences, as determined by comparing the sequences. In the art, identity also means the degree of sequence relatedness between polynucleotide sequences, as the case may be, as determined by the match between strings of such sequences. Both identity and similarity can be readily calculated (Lesk, A. M., 1988; Smith, D. W., 1993; Griffin, A. M., and Griffin, H. G., 1994; von Heinje, G., 1987; and Gribskov, M. and Devereux, J., 1991). While there exist a number of methods to measure identity and similarity between two polynucleotide sequences, both terms are well known to skilled artisans (von Heinje, G., 1987; Gribskov, M. and Devereux, 1991; and Carillo, H., and Lipman, D., 1988). Methods commonly employed to determine identity or similarity between sequences include, but are not limited to those disclosed in Carillo, H., and Lipman, D. (1988). Methods to determine identity and similarity are codified in computer programs. Computer program methods to determine identity and similarity between two sequences include, but are not limited to, GCG program package (Devereux, J., et al., 1984), BLASTP, BLASTP, and FASTA (Atschul, S. F. et al., 1990).

"Isolated" means altered "by the hand of man" from its natural state; i.e., that, if it occurs in nature, it has been changed or removed from its original environment, or both. For example, a naturally occurring polynucleotide naturally present in a living organism in its natural state is not "isolated," but the same polynucleotide separated from coexisting materials of its natural state is "isolated", as the term is employed herein. As part of or following isolation, such polynucleotides can be joined to other polynucleotides, such as DNA, for mutagenesis, to form fusion proteins, and for propagation or expression in a host, for instance. The isolated polynucleotides, alone or joined to other polynucleotides such as vectors, can be introduced into host cells, in culture or in whole organisms. Introduced into host cells in

culture or in whole organisms, such DNA still would be isolated, as the term is used herein, because they would not be in their naturally occurring form or environment. Similarly, the polynucleotides may occur in a composition, such as a media formulations, solutions for introduction of polynucleotides, for example, into cells, compositions or solutions for chemical or enzymatic reactions, for instance, which are not naturally occurring compositions, and, therein remain isolated polynucleotides within the meaning of that term as it is employed herein.

"Plasmids". Starting plasmids disclosed herein are either commercially available, publicly available, or can be constructed from available plasmids by routine application of well known, published procedures. Many plasmids and other cloning and expression vectors that can be used in accordance with the present invention are well known and readily available to those of skill in the art. Moreover, those of skill readily may construct any number of other plasmids suitable for use in the invention.

"Polynucleotides(s)" of the present invention may be in the form of RNA, such as mRNA, or in the form of DNA, including, for instance, cDNA and genomic DNA obtained by cloning or produced by chemical synthetic techniques or by a combination thereof. The DNA may be double-stranded or single-stranded. Single-stranded polynucleotides may be the coding strand, also known as the sense strand, or it may be the non-coding strand, also referred to as the anti-sense strand. Polynucleotides generally refers to any polyribonucleotide or polydeoxribonucleotide, which may be unmodified RNA or DNA or modified RNA or DNA. Thus, for instance, polynucleotides as used herein refers to, among others, single-and double-stranded DNA, DNA that is a mixture of single- and double-stranded regions or single-,

double- and triple-stranded regions, single- and double-stranded RNA, and RNA that is mixture of single- and double-stranded regions, hybrid molecules comprising DNA and RNA that may be single-stranded or, more typically, double-stranded, or triple-stranded, or a mixture of single- and double-stranded regions. In addition, polynucleotide as used herein refers to triple-stranded regions comprising RNA or DNA or both RNA and DNA. The strands in such regions may be from the same molecule or from different molecules. The regions may include all of one or more of the molecules, but more typically involve only a region of some of the molecules. One of the molecules of a triple-helical region often is an oligonucleotide. As used herein, the term polynucleotide also includes DNA or DNA that contain one or more modified bases. Thus, DNA or DNA with backbones modified for stability or for other reasons are "polynucleotides" as that term is intended herein. Moreover, DNA or DNA comprising unusual bases, such as inosine, or modified bases, such as tritylated bases, to name just two examples, are polynucleotides as the term is used herein. It will be appreciated that a great variety of modifications have been made to DNA and RNA that serve many useful purposes known to those of skill in the art. The term polynucleotide as it is employed herein embraces such chemically, enzymatically or metabolically modified forms of polynucleotides, as well as the chemical forms of DNA and RNA characteristic of viruses and cells, including simple and complex cells, inter alia. Polynucleotides embraces short polynucleotides often referred to as oligonucleotide(s). It will also be appreciated that RNA made by transcription of this doubled stranded nucleotide sequence, and an antisense strand of a nucleic acid molecule of the invention or an oligonucleotide fragment of the nucleic acid molecule, are contemplated within the scope of the invention. An antisense sequence is constructed by inverting the sequence of a nucleic acid molecule of the invention, relative to its normal presentation for transcription. Preferably, an antisense sequence is

constructed by inverting a region preceding the initiation codon or an unconserved region.

The antisense sequences may be constructed using chemical synthesis and enzymatic ligation reactions using procedures known in the art.

"Stringent hybridization conditions" are those which are stringent enough to provide specificity, reduce the number of mismatches and yet are sufficiently flexible to allow formation of stable hybrids at an acceptable rate. Such conditions are known to those skilled in the art and are described, for example, in Sambrook, et al, (1989). By way of example only, stringent hybridization with short nucleotides may be carried out at 5-10° below the T<sub>M</sub> using high concentrations of probe such as 0.01-1.0 pmole/ml. Preferably, the term "stringent conditions" means hybridization will occur only if there is at least 95% and preferably at least 97% identity between the sequences.

"Variant(s)" of polynucleotides are polynucleotides that differ in nucleotide sequence from another, reference polynucleotide. Generally, differences are limited so that the nucleotide sequences of the reference and the variant are closely similar overall and, in many regions, identical. Changes in the nucleotide sequence of the variant may be silent. That is, they may not alter the amino acids encoded by the polynucleotide. Where alterations are limited to silent changes of this type a variant will encode a polypeptide or polynucleotide with the same amino acid sequence as the reference. Changes in the nucleotide sequence of the variant may alter the amino acid sequence of a polypeptide encoded by the reference polynucleotide. Such nucleotide changes may result in amino acid substitutions, additions, deletions, fusions and truncations in the polypeptide or polynucleotide encoded by the reference sequence.

As hereinbefore mentioned, the present inventors have identified and sequenced a DNA sequence encoding PDE10A. The DNA sequence is shown in the Sequence Listing as SEQ ID NO:1, NO:2 and NO:11.

It will be appreciated that the invention includes nucleotide or amino acid sequences which have substantial sequence homology with the nucleotide sequences shown in the Sequence Listing as SEQ ID NO:1, NO:2 and NO:11. The term "sequences having substantial sequence homology" means those nucleotide and amino acid sequences which have slight or inconsequential sequence variations from the sequences disclosed in the Sequence Listing as SEQ ID NO:1, NO:2 and NO:11; i.e. the homologous sequences function in substantially the same manner to produce substantially the same polypeptides as the actual sequences. The variations may be attributable to local mutations or structural modifications. It is expected that a sequence having 85-90% sequence homology with the DNA sequence of the invention will provide a functional PDE10 polypeptide.

As used herein, "PDE10A" comprises a polynucleotide sequence which is down regulated in the course of CAG repeat disorders selected from the group consisting of: (a) a sequence comprising SEQ ID NO:1; (b) a sequence comprising SEQ ID NO:2; (c) a sequence comprising SEQ ID NO:11; (d) a sequence comprising nucleotides 257 to 2596 of SEQ ID NO:11; (e) a sequence which is at least 90% homologous with a sequence of (a), (b), (c) or (d), and; (f) a sequence which hybridizes to (a), (b), (c) or (d) under stringent conditions. In an embodiment, the isolated polynucleotide segment is cDNA. The invention also teaches an isolated polynucleotide segment, which retains substantially the same biological function or

activity as the polynucleotide encoded by the polynucleotide sequence.

Further embodiments of the invention are polynucleotides that are at least 70% identical over their entire length to a polynucleotide encoding PDE10 polypeptide or polynucleotide, and polynucleotides which are complementary to such polynucleotides. Other embodiments are polynucleotides that comprise a region that is at least 80% identical over their entire length to a polynucleotide encoding PDE10 of SEQ ID NO.11 and polynucleotides complementary thereto. This includes polynucleotides at least 90% identical over their entire length to the same, and among these embodiments are polynucleotides with at least 95%. Furthermore, those with at least 97% are highly preferred among those with at least 95%, and among these those with at least 98% and at least 99% are particularly highly preferred, with at least 99% being the more preferred.

The polynucleotides of the present invention may be employed as research reagents and materials for discovery of treatments of and diagnostics for disease, particularly human disease, as further discussed herein.

Analysis of the complete nucleotide and amino acid sequences of the protein of the invention using the procedures of Sambrook et al., supra, have been used to determine the expressed region, initiation codon and untranslated sequences of the PDE10A gene. The transcription regulatory sequences of the gene are determined by analyzing fragments of the DNA for their ability to express a reporter gene such as the bacterial gene lacZ.

The nucleic acid molecules of the invention allow those skilled in the art to construct

nucleotide probes for use in the detection of nucleotide sequences in biological materials. As shown in FIG. 11, 13, 15 and 16, a number of unique restriction sequences for restriction enzymes are incorporated in the nucleic acid molecule identified in the Sequence Listing as SEQ ID NO:1, NO:2 and NO:11, and these provide access to nucleotide sequences which code for polypeptides unique to the PDE10A polypeptide of the invention. Nucleotide sequences unique to PDE10A or isoforms thereof, can also be constructed by chemical synthesis and enzymatic ligation reactions carried out by procedures known in the art.

A nucleotide probe may be labeled with a detectable marker such as a radioactive label which provides for an adequate signal and has sufficient half-life such as 32p, 3H, 14C or the like. Other detectable markers which may be used include antigens that are recognized by a specific labeled antibody, fluorescent compounds, enzymes, antibodies specific for a labeled antigen, and chemiluminescent compounds. An appropriate label may be selected having regard to the rate of hybridization and binding of the probe to the nucleotide to be detected and the amount of nucleotide available for hybridization. The nucleotide probes may be used to detect genes related to or analogous to PDE10A of the invention.

Accordingly, the present invention also provides a method of detecting the presence of nucleic acid molecules encoding a polypeptide related to or analogous to PDE10A in a sample comprising contacting the sample under hybridization conditions with one or more of the nucleotide probes of the invention labeled with a detectable marker, and determining the degree of hybridization between the nucleic acid molecule in the sample and the nucleotide probes.

Hybridization conditions which may be used in the method of the invention are known in the art and are described for example in Sambrook J, et al., *supra*. The hybridization product may be assayed using techniques known in the art. The nucleotide probe may be labeled with a detectable marker as described herein and the hybridization product may be assayed by detecting the detectable marker or the detectable change produced by the detectable marker.

The nucleic acid molecule of the invention also permits the identification and isolation, or synthesis of nucleotide sequences which may be used as primers to amplify a polynucleotide molecule of the invention, for example in polymerase chain reaction (PCR). The length and bases of the primers for use in the PCR are selected so that they will hybridize to different strands of the desired sequence and at relative positions along the sequence such that an extension product synthesized from one primer when it is separated from its template can serve as a template for extension of the other primer into a nucleic acid of defined length.

Primers which may be used in the invention are oligonucleotides i.e. molecules containing two or more deoxyribonucleotides of the nucleic acid molecule of the invention which occur naturally as in a purified restriction endonuclease digest or are produced synthetically using techniques known in the art such as, for example, phosphotriester and phosphodiester methods (See Good et al, 1977) or automated techniques (see, for example, Conolly, B. A., 1987). The primers are capable of acting as a point of initiation of synthesis when placed under conditions which permit the synthesis of a primer extension product which is complementary to the DNA sequence of the invention e.g. in the presence of nucleotide substrates, an agent for polymerization such as DNA polymerase and at suitable temperature and pH. Preferably, the primers are sequences that do not form secondary structures by base

pairing with other copies of the primer or sequences that form a hair pin configuration. The primer may be single or double-stranded. When the primer is double-stranded it may be treated to separate its strands before using it to prepare amplification products. The primer preferably contains between about 7 and 25 nucleotides.

The primers may be labeled with detectable markers which allow for detection of the amplified products. Suitable detectable markers are radioactive markers such as P-32, S-35, I-125, and H-3, luminescent markers such as chemiluminescent markers, preferably luminol, and fluorescent markers, preferably dansyl chloride, fluorcein-5-isothiocyanate, and 4-fluor-7-nitrobenz-2-axa-1,3 diazole, enzyme markers such as horseradish peroxidase, alkaline phosphatase, .beta.-galactosidase, acetylcholinesterase, or biotin.

It will be appreciated that the primers may contain non-complementary sequences provided that a sufficient amount of the primer contains a sequence which is complementary to a nucleic acid molecule of the invention or oligonucleotide sequence thereof, which is to be amplified. Restriction site linkers may also be incorporated into the primers allowing for digestion of the amplified products with the appropriate restriction enzymes facilitating cloning and sequencing of the amplified product.

Thus, a method of determining the presence of a nucleic acid molecule having a sequence encoding PDE10A or a predetermined oligonucleotide fragment thereof in a sample, is provided comprising treating the sample with primers which are capable of amplifying the nucleic acid molecule or the predetermined oligonucleotide fragment thereof in a polymerase chain reaction to form amplified sequences, under conditions which permit the formation of

amplified sequences and, assaying for amplified sequences.

The polymerase chain reaction refers to a process for amplifying a target nucleic acid sequence as generally described in Innis et al, Academic Press, 1989, in Mullis et al., U.S. Pat. No. 4,863,195 and Mullis, U.S. Pat. No. 4,683,202 which are incorporated herein by reference. Conditions for amplifying a nucleic acid template are described in M. A. Innis and D. H. Gelfand, 1989, which is also incorporated herein by reference.

The amplified products can be isolated and distinguished based on their respective sizes using techniques known in the art. For example, after amplification, the DNA sample can be separated on an agarose gel and visualized, after staining with ethidium bromide, under ultra violet (UV) light. DNA may be amplified to a desired level and a further extension reaction may be performed to incorporate nucleotide derivatives having detectable markers such as radioactive labeled or biotin labeled nucleoside triphosphates. The primers may also be labeled with detectable markers. The detectable markers may be analyzed by restriction and electrophoretic separation or other techniques known in the art.

The conditions which may be employed in the methods of the invention using PCR are those which permit hybridization and amplification reactions to proceed in the presence of DNA in a sample and appropriate complementary hybridization primers. Conditions suitable for the polymerase chain reaction are generally known in the art. For example, see M. A. Innis and D. H. Gelfand, 1989, which is incorporated herein by reference. Preferably, the PCR utilizes polymerase obtained from the thermophilic bacterium Thermus aquatics (Taq polymerase, GeneAmp Kit, Perkin Elmer Cetus) or other thermostable polymerase may be used to amplify

DNA template strands.

It will be appreciated that other techniques such as the Ligase Chain Reaction (LCR) and Nucleic-Acid Sequence Based Amplification (NASBA) may be used to amplify a nucleic acid molecule of the invention. In LCR, two primers which hybridize adjacent to each other on the target strand are ligated in the presence of the target strand to produce a complementary strand (Barney, 1991 and European Published Application No. 0320308, published Jun. 14, 1989). NASBA is a continuous amplification method using two primers, one incorporating a promoter sequence recognized by an RNA polymerase and the second derived from the complementary sequence of the target sequence to the first primer (U.S. Ser. No. 5,130,238 to Malek).

The present invention also teaches vectors which comprise a polynucleotide or polynucleotides of the present invention, host cells which are genetically engineered with vectors of the invention and the production of polynucleotides of the invention by recombinant techniques.

In accordance with this aspect of the invention the vector may be, for example, a plasmid vector, a single or double-stranded phage vector, a single or double-stranded RNA or DNA viral vector. In certain embodiments in this regard, the vectors provide for specific expression. Such specific expression may be inducible expression or expression only in certain types of cells or both inducible and cell-specific. Particular among inducible vectors are vectors that can be induced for expression by environmental factors that are easy to manipulate, such as temperature and nutrient additives. A variety of vectors suitable to this

aspect of the invention, including constitutive and inducible expression vectors for use in prokaryotic and eukaryotic hosts, are well known and employed routinely by those of skill in the art. Such vectors include, among others, chromosomal, episomal and virus-derived vectors, e.g., vectors derived from bacterial plasmids, from bacteriophage, from transposons, from yeast episomes, from insertion elements, from yeast chromosomal elements, from viruses such as baculoviruses, papova viruses, such as SV40, vaccinia viruses, adenoviruses, fowl pox viruses, pseudorabies viruses and retroviruses, and vectors derived from combinations thereof, such as those derived from plasmid and bacteriophage genetic elements, such as cosmids and phagemids, all may be used for expression in accordance with this aspect of the present invention.

The following vectors, which are commercially available, are provided by way of example. Among vectors for use in bacteria are pQE70, pQE60 and pQE-9, available from Qiagen; pBS vectors, Phagescript vectors, Bluescript vectors, pNH8A, pNH16a, pNH18A, pNH46A, available from Stratagene; and ptrc99a, pKK223-3, pKK233-3, pDR540, pRIT5 available from Pharmacia, and pBR322 (ATCC 37017). Among eukaryotic vectors are pWLNEO, pSV2CAT, pOG44, pXT1 and pSG available from Stratagene; and pSVK3, pBPV, pMSG and pSVL available from Pharmacia. These vectors are listed solely by way of illustration of the many commercially available and well known vectors that are available to those of skill in the art for use in accordance with this aspect of the present invention. It will be appreciated that any other plasmid or vector suitable for, for example, introduction, maintenance, propagation or expression of a polynucleotide or polypeptide of the invention in a host may be used in this aspect of the invention. Generally, any vector suitable to maintain, propagate or express polynucleotides to express a polypeptide or polynucleotide in a host may be used

for expression in this regard.

The appropriate DNA sequence may be inserted into the vector by any of a variety of well-known and routine techniques. In general, expression constructs will contain sites for transcription initiation and termination, and, in the transcribed region, a ribosome binding site for translation. The coding portion of the mature transcripts expressed by the constructs will include a translation initiating AUG at the beginning and a termination codon appropriately positioned at the end of the polynucleotide to be translated.

The DNA sequence in the expression vector is operatively linked to appropriate expression control sequence(s), including, for instance, a promoter to direct mRNA transcription. Promoter regions can be selected from any desired gene using vectors that contain a reporter transcription unit lacking a promoter region, such as a chloramphenicol acetyl transferase ("CAT") transcription unit, downstream of restriction site or sites for introducing a candidate promoter fragment; i.e., a fragment that may contain a promoter. As is well known, introduction into the vector of a promoter-containing fragment at the restriction site upstream of the cat gene engenders production of CAT activity, which can be detected by standard CAT assays. Vectors suitable to this end are well known and readily available, such as pKK232-8 and pCM7. Promoters for expression of polynucleotides of the present invention include not only well known and readily available promoters, but also promoters that readily may be obtained by the foregoing technique, using a reporter gene. Among known prokaryotic promoters suitable for expression of polynucleotides and polypeptides in accordance with the present invention are the E. coli lacI and lacZ and promoters, the T3 and T7 promoters, the gpt promoter, the lambda PR, PL promoters and the trp promoter. Among

known eukaryotic promoters suitable in this regard are the CMV immediate early promoter, the HSV thymidine kinase promoter, the early and late SV40 promoters, the promoters of retroviral LTRs, such as those of the Rous sarcoma virus ("RSV"), and metallothionein promoters, such as the mouse metallothionein-I promoter.

Vectors for propagation and expression generally will include selectable markers and amplification regions, such as, for example, those set forth in Sambrook et al., supra.

As hereinbefore mentioned, the present invention also teaches host cells which are genetically engineered with vectors of the invention.

Polynucleotide constructs in host cells can be used in a conventional manner to produce the gene product encoded by the recombinant sequence. The PDE10A polynucleotide or polypeptide products or isoforms or parts thereof, may be obtained by expression in a suitable host cell using techniques known in the art. Suitable host cells include prokaryotic or eukaryotic organisms or cell lines, for example bacterial, mammalian, yeast, or other fungi, viral, plant or insect cells. Methods for transforming or transfecting cells to express foreign DNA are well known in the art (See for example, Itakura et al., U.S. Pat. No. 4,704,362; Hinnen et al., 1978; Murray et al., U.S. Pat. No. 4,801,542; Upshall et al., U.S. Pat. No. 4,935,349; Hagen et al., U.S. Pat. No. 4,784,950; Axel et al., U.S. Pat. No. 4,399,216; Goeddal et al., U.S. Pat. No. 4,766,075; and Sambrook et al, 1989, all of which are incorporated herein by reference). Representative examples of appropriate hosts include bacterial cells, such as streptococci, staphylococci, E. coli, streptomyces and Bacillus subtilis cells; fungal cells, such as yeast cells and Aspergillus cells; insect cells such as Drosophila S2

and Spodoptera Sf9 cells; animal cells such as CHO, COS, HeLa, C127, 3T3, BHK, 293 and Bowes melanoma cells; and plant cells.

Host cells can be genetically engineered to incorporate polynucleotides and express polynucleotides of the present invention. Introduction of polynucleotides into the host cell can be affected by calcium phosphate transfection, DEAE-dextran mediated transfection, transvection, microinjection, cationic lipid-mediated transfection, electroporation, transduction, scrape loading, ballistic introduction, infection or other methods. Such methods are described in many standard laboratory manuals, such as Davis et al. (1986) and Sambrook et al. (1989).

As hereinbefore mentioned, the present invention also teaches the production of polynucleotides of the invention by recombinant techniques.

The PDE10 polynucleotides encode a polypeptide which is the mature protein plus additional amino or carboxyl-terminal amino acids, or amino acids interior to the mature polypeptide (when the mature form has more than one polypeptide chain, for instance). Such sequences may play a role in processing of a protein from precursor to a mature form, may allow protein transport, may lengthen or shorten protein half-life or may facilitate manipulation of a protein for assay or production, among other things. As generally is the case in vivo, the additional amino acids may be processed away from the mature protein by cellular enzymes.

A precursor protein, having the mature form of the polypeptide fused to one or more prosequences may be an inactive form of the polypeptide. When prosequences are removed

such inactive precursors generally are activated. Some or all of the prosequences may be removed before activation. Generally, such precursors are called proproteins.

In sum, a polynucleotide of the present invention may encode a mature protein, a mature protein plus a leader sequence (which may be referred to as a preprotein), a precursor of a mature protein having one or more prosequences which are not the leader sequences of a preprotein, or a preproprotein, which is a precursor to a proprotein, having a leader sequence and one or more prosequences, which generally are removed during processing steps that produce active and mature forms of the polypeptide.

The polypeptides of the invention may be prepared by culturing the host/vector systems described above, in order to express the recombinant polypeptides. Recombinantly produced PDE10A based protein or parts thereof, may be further purified using techniques known in the art such as commercially available protein concentration systems, by salting out the protein followed by dialysis, by affinity chromatography, or using anion or cation exchange resins.

Mature proteins can be expressed in mammalian cells, yeast, bacteria, or other cells under the control of appropriate promoters. Cell-free translation systems can also be employed to produce such proteins using DNA derived from the DNA constructs of the present invention. Appropriate cloning and expression vectors for use with prokaryotic and eukaryotic hosts are described by Sambrook et al., supra.

Polynucleotides of the invention, encoding the heterologous structural sequence of a

polynucleotide or polypeptide of the invention generally will be inserted into a vector using standard techniques so that it is operably linked to the promoter for expression. The polynucleotide will be positioned so that the transcription start site is located appropriately 5' to a ribosome binding site. The ribosome binding site will be 5' to the AUG that initiates translation of the polynucleotide or polypeptide to be expressed. Generally, there will be no other open reading frames that begin with an initiation codon, usually AUG, and lie between the ribosome binding site and the initiation codon. Also, generally, there will be a translation stop codon at the end of the expressed polynucleotide and there will be a polyadenylation signal in constructs for use in eukaryotic hosts. Transcription termination signal appropriately disposed at the 3' end of the transcribed region may also be included in the polynucleotide construct.

For secretion of the translated protein into the lumen of the endoplasmic reticulum, into the periplasmic space or into the extracellular environment, appropriate secretion signals may be incorporated into the expressed polynucleotide or polypeptide. These signals may be endogenous to the polynucleotide or they may be heterologous signals. Microbial cells employed in expression of proteins can be disrupted by any convenient method, including freeze-thaw cycling, sonication, mechanical disruption, or use of cell lysing agents, such methods are well know to those skilled in the art. PDE10A polynucleotide or polypeptide can be recovered and purified from recombinant cell cultures by well-known methods including ammonium sulfate or ethanol precipitation, acid extraction, anion or cation exchange chromatography, phosphocellulose chromatography, hydrophobic interaction chromatography, affinity chromatography, hydroxylapatite chromatography and lectin chromatography. Most preferably, high performance liquid chromatography is employed for

purification. Well known techniques for refolding protein may be employed to regenerate active conformation when the polynucleotide is denatured during isolation and or purification.

In an embodiment, a nucleic acid molecule of the invention may be cloned into a glutathione S-transferase (GST) gene fusion system for example the pGEX-1 T, pGEX-2T and pGEX-3X of Pharmacia. The fused gene may contain a strong lac promoter, inducible to a high level of expression by IPTG, as a regulatory element. Thrombin or factor Xa cleavage sites may be present which allow proteolytic cleavage of the desired polypeptide from the fusion product. The glutathione S-transferase-PDE10A fusion protein may be easily purified using a glutathione sepharose 4B column, for example from Pharmacia. The 26 kd glutathione S-transferase polypeptide can be cleaved by thrombin (pGEX-1 or pGEX-2T) or factor Xa (pGEX-3X) and resolved from the using the polypeptide using the same affinity column. Additional chromatographic steps can be included if necessary, for example Sephadex or DEAE cellulose. The two enzymes may be monitored by protein and enzymatic assays and purity may be confirmed using SDS-PAGE.

The PDE10A protein or parts thereof may also be prepared by chemical synthesis using techniques well known in the chemistry of proteins such as solid phase synthesis (Merrifield, 1964) or synthesis in homogenous solution (Houbenweyl, 1987).

Within the context of the present invention, PDE10A polypeptide includes various structural forms of the primary protein which retain biological activity. For example, PDE10A polypeptide may be in the form of acidic or basic salts or in neutral form. In addition,

individual amino acid residues may be modified by oxidation or reduction. Furthermore, various substitutions, deletions or additions may be made to the amino acid or nucleic acid sequences, the net effect being that biological activity of PDE10A is retained. Due to code degeneracy, for example, there may be considerable variation in nucleotide sequences encoding the same amino acid.

The polypeptide may be expressed in a modified form, such as a fusion protein, and may include not only secretion signals but also additional heterologous functional regions. Thus, for instance, a region of additional amino acids, particularly charged amino acids, may be added to the C- or N-terminus of the polypeptide to improve stability and persistence in the host cell, during purification or during subsequent handling and storage. Also, fusion proteins may be added to the polynucleotide or polypeptide to facilitate purification. Such regions may be removed prior to final preparation of the polynucleotide or polypeptide. The addition of peptide moieties to polynucleotide or polypeptides to engender secretion or excretion, to improve stability or to facilitate purification, among others, are familiar and routine techniques in the art. In drug discovery, for example, proteins have been fused with antibody Fc portions for the purpose of high-throughput screening assays to identify antagonists (see Bennett et al., 1995, and Johanson et al., 1995).

Detecting Presence of or Predisposition for CAG Repeat Disorders

This invention is also related to the use of the PDE10A polynucleotides to detect complementary polynucleotides as a diagnostic reagent. Detection of the level of expression of PDE10A in a eukaryote, particularly a mammal, and especially a human, will provide a

method for diagnosis of a disease. Eukaryotes (herein also "individual(s)"), particularly mammals, and especially humans, exhibiting decreased levels of PDE10A may be detected by a variety of techniques. Nucleic acids for diagnosis may be obtained from an infected individual's cells and tissues, such as the striatum, nucleus accumbens and olfactory tubercule. RNA may be used directly for detection or may be amplified enzymatically by using PCR (Saiki et al., 1986) prior to analysis. As an example, PCR primers complementary to the nucleic acid encoding PDE10A can be used to identify and analyze PDE10A presence and/or expression. Using PCR, characterization of the level of PDE10A present in the individual may be made by comparative analysis.

The invention thus provides a process for detecting disease by using methods known in the art and methods described herein to detect decreased expression of PDE10 polynucleotide. For example, decreased expression of PDE10 polynucleotide can be measured using any on of the methods well known in the art for the quantification of polynucleotides, such as, for example, PCR, RT-PCR, DNAse protection, northern blotting and other hybridization methods. Thus, the present invention provides a method for detecting triplet-repeat disorders, and a method for detecting a genetic pre-disposition for triplet-repeat disorders and other disorders of the basal ganglia including schizophrenia, stroke, trauma, Parkinson's disease and Alzheimer's disease (AD). More generally, the present invention provides a method for detecting a genetic pre-disposition for neurological disorders characterized by progressive cell loss.

Drug Screening Assays

The invention also provides a method of screening compounds to identify those which enhance (agonist) or block (antagonist) the action of PDE10 polypeptides or polynucleotides, such as its interaction with PDE10-binding molecules. The identification of mutations in specific genes in inherited neurodegenerative disorders, combined with advances in the field of transgenic methods, provides those of skill in the art with the information necessary to further study human diseases. This is extraordinarily useful in modeling familial forms of triplet-repeat disorders and other disorders of the basal ganglia including schizophrenia, stroke, trauma, Parkinson's disease and Alzheimer's disease (AD). More generally, the present invention is useful for modeling neurological disorders characterized by progressive cell loss, as well as those involving acute cell loss, such as stroke and trauma.

For example, to screen for agonists or antagonists, a synthetic reaction mix, a cellular compartment, such as a membrane, cell envelope or cell wall, or a preparation of any thereof, may be prepared from a cell that expresses a molecule that binds PDE10. The preparation is incubated with labeled PDE10 in the absence or the presence of a candidate molecule which may be a PDE10 agonist or antagonist. The ability of the candidate molecule to bind the binding molecule is reflected in decreased binding of the labeled ligand.

PDE10-like effects of potential agonists and antagonists may by measured, for instance, by determining activity of a reporter system following interaction of the candidate molecule with a cell or appropriate cell preparation, and comparing the effect with that of PDE10 or molecules that elicit the same effects as PDE10. Reporter systems that may be useful in this

regard include, but are not limited to, colorimetric labeled substrate converted into product, a reporter gene that is responsive to changes in PDE10 activity, and binding assays known in the art.

Another example of an assay for PDE10 antagonists is a competitive assay that combines PDE10 and a potential antagonist with membrane-bound PDE10-binding molecules, recombinant PDE10 binding molecules, natural substrates or ligands, or substrate or ligand mimetics, under appropriate conditions for a competitive inhibition assay. PDE10 can be labeled, such as by radioactivity or a colorimetric compound, such that the number of PDE10 molecules bound to a binding molecule or converted to product can be determined accurately to assess the effectiveness of the potential antagonist.

Potential antagonists include small organic molecules, peptides, polypeptides and antibodies that bind to a polynucleotide or polypeptide of the invention and thereby inhibit or extinguish its activity. Potential antagonists also may be small organic molecules, a peptide, a polypeptide such as a closely related protein or antibody that binds the same sites on a binding molecule, such as a binding molecule, without inducing PDE10-induced activities, thereby preventing the action of PDE10 by excluding PDE10 from binding.

Potential antagonists include a small molecule which binds to and occupies the binding site of the polypeptide thereby preventing binding to cellular binding molecules, such that normal biological activity is prevented. Examples of small molecules include but are not limited to small organic molecules, peptides or peptide-like molecules. Other potential antagonists include antisense molecules (see Okano, 1988, for a description of these molecules).

Potential antagonists include compounds related to and derivatives of PDE10.

Developing modulators of the biological activities of specific PDEs requires differentiating PDE isozymes present in a particular assay preparation. The classical enzymological approach of isolating PDEs from natural tissue sources and studying each new isozyme may be used. Another approach has been to identify assay conditions which might favor the contribution of one isozyme and minimize the contribution of others in a preparation. Still another approach has been the separation of PDEs by immunological means. Each of the foregoing approaches for differentiating PDE isozymes is time consuming. As a result many attempts to develop selective PDE modulators have been performed with preparations containing more than one isozyme. Moreover, PDE preparations from natural tissue sources are susceptible to limited proteolysis and may contain mixtures of active proteolytic products that have different kinetic, regulatory and physiological properties than the full length PDEs.

Recombinant PDE10 polypeptide products of the invention greatly facilitate the development of new and specific PDE10 modulators. The need for purification of an isozyme can be avoided by expressing it recombinantly in a host cell that lacks endogenous phosphodiesterase activity (e.g., yeast strain YKS45 deposited as ATCC 74225). Once a compound that modulates the activity of the PDE10 is discovered, its selectivity can be evaluated by comparing its activity on the PDE10 to its activity on other PDE isozymes. Thus, the combination of the recombinant PDE10 products of the invention with other recombinant PDE products in a series of independent assays provides a system for developing selective modulators of PDE10. Selective modulators may include, for example, antibodies and other proteins or peptides which specifically bind to the PDE10 or PDE10 nucleic acid,

International Publication No. WO93/05182 published Mar. 18, 1993 which describes methods for selecting oligonucleotides which selectively bind to target biomolecules) or PDE10 nucleic acid (e.g., antisense oligonucleotides) and other non-peptide natural or synthetic compounds which specifically bind to the PDE10 or PDE10 nucleic acid. Mutant forms of the PDE10 which alter the enzymatic activity of the PDE10 or its localization in a cell are also contemplated. Crystallization of recombinant PDE10 alone and bound to a modulator, analysis of atomic structure by X-ray crystallography, and computer modelling of those structures are methods useful for designing and optimizing non-peptide selective modulators. See, for example, Erickson et al., *Ann. Rep. Med. Chem.*, 27: 271-289 (1992) for a general review of structure-based drug design.

Targets for the development of selective modulators include, for example: (1) the regions of the PDE10 which contact other proteins and/or localize the PDE10 within a cell, (2) the regions of the PDE10 which bind substrate, (3) the allosteric cGMP-binding site(s) of PDE10, (4) the metal-binding regions of the PDE10, (5) the phosphorylation site(s) of PDE10 and (6) the regions of the PDE10 which are involved in dimerization of PDE10 subunits.

Thus, the present invention provides a method for screening and selecting compounds which promote triplet-repeat disorders, and a method for screening and selecting compounds which treat or inhibit triplet-repeat disorders, as well as schizophrenia, stroke, trauma, Parkinson's disease and Alzheimer's disease. More generally, the present invention provides a method for screening and selecting compounds which promote or inhibit neurological disorders characterized by progressive cell loss, as well as those involving acute cell loss, such as

stroke and trauma.

The selected antagonists and agonists may be administered, for instance, to inhibit progressive and acute neurological disorders, such as Huntington's disease, Parkinson's disease, schizophrenia, Alzheimer's disease (AD), stroke or trauma.

Antagonists and agonists and other compounds of the present invention may be employed alone or in conjunction with other compounds, such as therapeutic compounds. The pharmaceutical compositions may be administered in any effective, convenient manner including, for instance, administration by direct microinjection into the affected area, or by intravenous or other routes. These compositions of the present invention may be employed in combination with a non-sterile or sterile carrier or carriers for use with cells, tissues or organisms, such as a pharmaceutical carrier suitable for administration to a subject. Such compositions comprise, for instance, a media additive or a therapeutically effective amount of antagonists or agonists of the invention and a pharmaceutically acceptable carrier or excipient. Such carriers may include, but are not limited to, saline, buffered saline, dextrose, water, glycerol, ethanol and combinations thereof. The formulation is prepared to suit the mode of administration.

Inhibition of PDE10A will be highly detrimental to striatal brain function. The progressive decline in PDE10A mRNA levels in HD may lead to dysregulation of cAMP levels and neuronal dysfunction. Up-regulation of PDE10A will be effective in combating such neuronal dysfunction.

Gene Therapy

A variety of gene therapy approaches may be used in accordance with the invention to modulate expression of the PDE10A gene in vivo. For example, antisense DNA molecules may be engineered and used to block translation of PDE10A mRNA in vivo. Alternatively, ribozyme molecules may be designed to cleave and destroy the PDE10A mRNAs in vivo. In another alternative, oligonucleotides designed to hybridize to the 5' region of the PDE10A gene (including the region upstream of the coding sequence) and form triple helix structures may be used to block or reduce transcription of the PDE10A gene. In yet another alternative, nucleic acid encoding the full length wild-type PDE10A message may be introduced in vivo into cells which otherwise would be unable to produce the wild-type PDE10A gene product in sufficient quantities or at all.

In a preferred embodiment, the antisense, ribozyme and triple helix nucleotides are designed to inhibit the translation or transcription of PDE10A. To accomplish this, the oligonucleotides used should be designed on the basis of relevant sequences unique to PDE10A.

For example, and not by way of limitation, the oligonucleotides should not fall within those region where the nucleotide sequence of PDE10A is most homologous to that of other PDEs, such as PDE2 PDE5 and PDE6, herein referred to as "unique regions".

In the case of antisense molecules, it is preferred that the sequence be chosen from the unique regions. It is also preferred that the sequence be at least 18 nucleotides in length in order to

achieve sufficiently strong annealing to the target mRNA sequence to prevent translation of the sequence. Izant and Weintraub, 1984, Cell, 36:1007-1015; Rosenberg et al., 1985, Nature, 313:703-706.

In the case of the "hammerhead" type of ribozymes, it is also preferred that the target sequences of the ribozymes be chosen from the unique regions. Ribozymes are RNA molecules which possess highly specific endoribonuclease activity. Hammerhead ribozymes comprise a hybridizing region which is complementary in nucleotide sequence to at least part of the target RNA, and a catalytic region which is adapted to cleave the target RNA. The hybridizing region contains nine (9) or more nucleotides. Therefore, the hammerhead ribozymes of the present invention have a hybridizing region which is complementary to the sequences listed above and is at least nine nucleotides in length. The construction and production of such ribozymes is well known in the art and is described more fully in Haseloff and Gerlach, 1988, Nature, 334:585-591.

The ribozymes of the present invention also include RNA endoribonucleases (hereinafter "Cech-type ribozymes") such as the one which occurs naturally in Tetrahymena Thermophila (known as the IVS, or L-19 IVS RNA) and which has been extensively described by Thomas Cech and collaborators (Zaug, et al., 1984, Science, 224:574-578; Zaug and Cech, 1986, Science, 231:470-475; Zaug, et al., 1986, Nature, 324:429-433; published International patent application No. WO 88/04300 by University Patents Inc.; Been and Cech, 1986, Cell, 47:207-216). The Cech endoribonucleases have an eight base pair active site which hybridizes to a target RNA sequence whereafter cleavage of the target RNA takes place. The invention encompasses those Cech-type ribozymes which target eight base-pair active site sequences

that are present in PDE10A but not other PDEs.

The foregoing compounds can be administered by a variety of methods which are known in the art including, but not limited to the use of liposomes as a delivery vehicle. Naked DNA or RNA molecules may also be used where they are in a form which is resistant to degradation such as by modification of the ends, by the formation of circular molecules, or by the use of alternate bonds including phosphothionate and thiophosphoryl modified bonds. In addition, the delivery of nucleic acid may be by facilitated transport where the nucleic acid molecules are conjugated to poly-lysine or transferrin. Nucleic acid may also be transported into cells by any of the various viral carriers, including but not limited to, retrovirus, vaccinia, AAV, and adenovirus.

Alternatively, a recombinant nucleic acid molecule which encodes, or is, such antisense, ribozyme, triple helix, or PDE10A molecule can be constructed. This nucleic acid molecule may be either RNA or DNA. If the nucleic acid encodes an RNA, it is preferred that the sequence be operatively attached to a regulatory element so that sufficient copies of the desired RNA product are produced. The regulatory element may permit either constitutive or regulated transcription of the sequence. In vivo, that is, within the cells or cells of an organism, a transfer vector such as a bacterial plasmid or viral RNA or DNA, encoding one or more of the RNAs, may be transfected into cells e.g. (Llewellyn et al., 1987, J. *Mol. Biol.*, 195:115-123; Hanahan et al. 1983, *J. Mol. Biol.*, 166:557-580). Once inside the cell, the transfer vector may replicate, and be transcribed by cellular polymerases to produce the RNA or it may be integrated into the genome of the host cell. Alternatively, a transfer vector containing sequences encoding one or more of the RNAs may be transfected into cells or

introduced into cells by way of micromanipulation techniques such as microinjection, such that the transfer vector or a part thereof becomes integrated into the genome of the host cell.

Composition, Formulation, and Administration of Pharmaceutical Compositions

The pharmaceutical compositions of the present invention may be manufactured in a manner that is itself known, e.g., by means of conventional mixing, dissolving, granulating, dragee-making, levigating, emulsifying, encapsulating, entrapping or lyophilizing processes.

Pharmaceutical compositions for use in accordance with the present invention thus may be formulated in conventional manner using one or more physiologically acceptable carriers comprising excipients and auxiliaries which facilitate processing of the active compounds into preparations which can be used pharmaceutically. Proper formulation is dependent upon the route of administration chosen.

For injection, the agents of the invention may be formulated in aqueous solutions, preferably in physiologically compatible buffers such as Hanks's solution, Ringer's solution, or physiological saline buffer. For transmucosal administration, penetrants appropriate to the barrier to be permeated are used in the formulation. Such penetrants are generally known in the art.

For oral administration, the compounds can be formulated readily by combining the active compounds with pharmaceutically acceptable carriers well known in the art. Such carriers enable the compounds of the invention to be formulated as tablets, pills, dragees, capsules,

liquids, gels, syrups, slurries, suspensions and the like, for oral ingestion by a patient to be treated. Pharmaceutical preparations for oral use can be obtained solid excipient, optionally grinding a resulting mixture, and processing the mixture of granules, after adding suitable auxiliaries, if desired, to obtain tablets or dragee cores. Suitable excipients are, in particular, fillers such as sugars, including lactose, sucrose, mannitol, or sorbitol; cellulose preparations such as, for example, maize starch, wheat starch, rice starch, potato starch, gelatin, gum tragacanth, methyl cellulose, hydroxypropylmethyl-cellulose, sodium carboxymethylcellulose, and/or polyvinylpyrrolidone (PVP). If desired, disintegrating agents may be added, such as the cross-linked polyvinyl pyrrolidone, agar, or alginic acid or a salt thereof such as sodium alginate.

Dragee cores are provided with suitable coatings. For this purpose, concentrated sugar solutions may be used, which may optionally contain gum arabic, talc, polyvinyl pyrrolidone, carbopol gel, polyethylene glycol, and/or titanium dioxide, lacquer solutions, and suitable organic solvents or solvent mixtures. Dyestuffs or pigments may be added to the tablets or dragee coatings for identification or to characterize different combinations of active compound doses.

Pharmaceutical preparations which can be used orally include push-fit capsules made of gelatin, as well as soft, sealed capsules made of gelatin and a plasticizer, such as glycerol or sorbitol. The push-fit capsules can contain the active ingredients in admixture with filler such as lactose, binders such as starches, and/or lubricants such as talc or magnesium stearate and, optionally, stabilizers. In soft capsules, the active compounds may be dissolved or suspended in suitable liquids, such as fatty oils, liquid paraffin, or liquid polyethylene glycols. In

addition, stabilizers may be added. All formulations for oral administration should be in dosages suitable for such administration.

For buccal administration, the compositions may take the form of tablets or lozenges formulated in conventional manner.

For administration by inhalation, the compounds for use according to the present invention are conveniently delivered in the form of an aerosol spray presentation from pressurized packs or a nebulizer, with the use of a suitable propellant, e.g., dichlorodifluoromethane, trichlorofluoromethane, dichlorotetrafluoroethane, carbon dioxide or other suitable gas. In the case of a pressurized aerosol the dosage unit may be determined by providing a valve to deliver a metered amount. Capsules and cartridges of e.g. gelatin for use in an inhaler or insufflator may be formulated containing a powder mix of the compound and a suitable powder base such as lactose or starch.

The compounds may be formulated for parenteral administration by injection, e.g., by bolus injection or continuous infusion. Formulations for injection may be presented in unit dosage form, e.g., in ampoules or in multidose containers, with an added preservative. The compositions may take such forms as suspensions, solutions or emulsions in oily or aqueous vehicles, and may contain formulatory agents such as suspending, stabilizing and/or dispersing agents.

Pharmaceutical formulations for parenteral administration include aqueous solutions of the active compounds in water-soluble form. Additionally, suspensions of the active compounds

may be prepared as appropriate oily injection suspensions. Suitable lipophilic solvents or vehicles include fatty oils such as sesame oil, or synthetic fatty acid esters, such as ethyl oleate or triglycerides, or liposomes. Aqueous injection suspensions may contain substances which increase the viscosity of the suspension, such as sodium carboxymethyl cellulose, sorbitol, or dextran. Optionally, the suspension may also contain suitable stabilizers or agents which increase the solubility of the compounds to allow for the preparation of highly concentrated solutions.

Alternatively, the active ingredient may be in powder form for constitution with a suitable vehicle, e.g., sterile pyrogen-free water, before use.

The compounds may also be formulated in rectal compositions such as suppositories or retention enemas, e.g., containing conventional suppository bases such as cocoa butter or other glycerides.

In addition to the formulations described previously, the compounds may also be formulated as a depot preparation. Such long acting formulations may be administered by implantation (for example subcutaneously or intramuscularly) or by intramuscular injection. Thus, for example, the compounds may be formulated with suitable polymeric or hydrophobic materials (for example as an emulsion in an acceptable oil) or ion exchange resins, or as sparingly soluble derivatives, for example, as a sparingly soluble salt.

A pharmaceutical carrier for the hydrophobic compounds of the invention is a cosolvent system comprising benzyl alcohol, a nonpolar surfactant, a water-miscible organic polymer,

and an aqueous phase. Naturally, the proportions of a co-solvent system may be varied considerably without destroying its solubility and toxicity characteristics. Furthermore, the identity of the co-solvent components may be varied.

Alternatively, other delivery systems for hydrophobic pharmaceutical compounds may be employed. Liposomes and emulsions are well known examples of delivery vehicles or carriers for hydrophobic drugs. Certain organic solvents such as dimethylsulfoxide also may be employed, although usually at the cost of greater toxicity. Additionally, the compounds may be delivered using a sustained-release system, such as semipermeable matrices of solid hydrophobic polymers containing the therapeutic agent. Various of sustained-release materials have been established and are well known by those skilled in the art. Sustained-release capsules may, depending on their chemical nature, release the compounds for a few weeks up to over 100 days. Depending on the chemical nature and the biological stability of the therapeutic reagent, additional strategies for protein stabilization may be employed.

The pharmaceutical compositions also may comprise suitable solid or gel phase carriers or excipients. Examples of such carriers or excipients include but are not limited to calcium carbonate, calcium phosphate, various sugars, starches, cellulose derivatives, gelatin, and polymers such as polyethylene glycols.

Many of the compounds of the invention may be provided as salts with pharmaceutically compatible counterions. Pharmaceutically compatible salts may be formed with many acids, including but not limited to hydrochloric, sulfuric, acetic, lactic, tartaric, malic, succinic, etc. Salts tend to be more soluble in aqueous or other protonic solvents that are the corresponding

free base forms.

Suitable routes of administration may, for example, include oral, rectal, transmucosal, transdermal, or intestinal administration; parenteral delivery, including intramuscular, subcutaneous, intramedullary injections, as well as intrathecal, direct intraventricular, intravenous, intraperitoneal, intranasal, or intraocular injections.

Alternately, one may administer the compound in a local rather than systemic manner, for example, via injection of the compound directly into an affected area, often in a depot or sustained release formulation.

Furthermore, one may administer the drug in a targeted drug delivery system, for example, in a liposome coated with an antibody specific for affected cells. The liposomes will be targeted to and taken up selectively by the cells.

The pharmaceutical compositions generally are administered in an amount effective for treatment or prophylaxis of a specific indication or indications. It is appreciated that optimum dosage will be determined by standard methods for each treatment modality and indication, taking into account the indication, its severity, route of administration, complicating conditions and the like. In therapy or as a prophylactic, the active agent may be administered to an individual as an injectable composition, for example as a sterile aqueous dispersion, preferably isotonic. A therapeutically effective dose further refers to that amount of the compound sufficient to result in amelioration of symptoms associated with such disorders. Techniques for formulation and administration of the compounds of the instant

application may be found in "Remington's Pharmaceutical Sciences," Mack Publishing Co., Easton, Pa., latest edition. For administration to mammals, and particularly humans, it is expected that the daily dosage level of the active agent will be from 0.001 mg/kg to 10 mg/kg, typically around 0.01 mg/kg. The physician in any event will determine the actual dosage which will be most suitable for an individual and will vary with the age, weight and response of the particular individual. The above dosages are exemplary of the average case. There can, of course, be individual instances where higher or lower dosage ranges are merited, and such are within the scope of this invention.

The invention further provides diagnostic and pharmaceutical packs and kits comprising one or more containers filled with one or more of the ingredients of the aforementioned compositions of the invention. Associated with such container(s) can be a notice in the form prescribed by a governmental agency regulating the manufacture, use or sale of pharmaceuticals or biological products, reflecting approval by the agency of the manufacture, use or sale of the product for human administration.

#### **EXAMPLES**

The present invention is further described by the following examples. These examples, while illustrating certain specific aspects of the invention, do not portray the limitations or circumscribe the scope of the disclosed invention.

EXAMPLE 1 - Isolation of PDE10A

Wild-type (B6CBAF1) and HD transgenic [B6CBA-TgN(Hdexon1)62Gpb] mice (Jackson Laboratories) and adult Sprague-Dawley rats (250-300 g; Charles River Laboratories) and were used in this study. The genotype of the mice was determined by PCR amplification of a 100 bp region of the integrated human HD exon 1 transgene using primers corresponding to nts 3340-3459 (5'-AGG GCT GTC AAT CAT GCT GG-3') and nts 3836-3855 (5'-AAA CTC ACG GTC GGT GCA GC-3') of clone E4.1 of the human HD gene (Accession number L34020). PCR conditions used are described in Mangiarini et al.(1996). DNA was extracted from a tail clip and an ear punch from each mouse used in this study. Both samples were subjected to PCR genotype analysis. For *in situ* hybridization analysis, the animals were anesthetized with >100 mg/kg sodium pentobarbital, decapitated, the brains removed and stored at -70°C prior to sectioning. For RNA isolation, animals were anesthetized, decapitated and the striatum and cortex were excised and stored in liquid nitrogen prior to RNA extraction. Animal care was given according to protocols approved by Dalhousie University and the Canadian Council of Animal Care.

Differential display was used to identify novel mDNA or previously described mDNA whose relative expression levels are altered as a result of the presence of the transgene. Using differential display, the mRNA populations derived from the striatum of 10 week old wild type were compared with age-matched R6/2 transgenic mice. Differential display has been used extensively (> 750 references) since its development (Liang and Pardee, 1992) to identify changes in gene expression in cells and in tissues including brain (Douglass et al., 1995; Babity et al., 1997a; Livesey et al., 1997; Berke et al., 1998). Perhaps the most

important finding was the demonstration by Qu et al., (1996) that differential display can be used to isolate genes differentially expressed in inbred strains of mice. The power of differential display is that the sequence information obtained can be directly related to the experimental paradigm. Moreover, such sequence information includes sufficient information to identify transcripts and can then lead to experiments that reveal function of the cognate protein in the experimental model.

DNA sequence information of potentially differentially expressed cDNA can be used to generate oligonucleotide probes for in situ hybridization to define the anatomical and temporal patterns of expression of specific transcripts (see Babity et al., 1997a). This technique is especially useful to study changes in steady-state levels of mRNA in heterogeneous tissue such as brain. Brain tissue can be micro-dissected (Babity et al., 1997b). This enabled the present inventors to reduce the requirement for tissue, and hence compare the mRNA populations derived from individual animals for each experimental group.

Thus RT-PCR (Denovan-Wright et al., 1999) was used to identify differences in the patterns of gene expression between the striatum of wild-type and transgenic mice that were hemizygous for the 5' UTR, exon 1 and part of intron 1 of the human Huntingon's Disease gene. Total cellular RNA was isolated from the striatum and cortex of three 10 week-old wild-type and three 10 week-old R6/2 HD mice (Mangiarini et al., 1996) and used as the template to generate single-stranded cDNA. Total cellular RNA from each animal and tissue was purified using Trizol™ reagent (Gibco BRL) and the manufacture's protocol. 10 µg aliquots of total RNA were treated with RQ1 DNAse-free DNAse (Promega) in the presence

of DNAsin<sup>TM</sup> (Promega) DNAse inhibitor to remove trace genomic DNA and then converted to single-stranded cDNA. The primers and conditions for PCR amplification follow those of the Delta<sup>TM</sup> RNA fingerprinting manual (Clontech).

The cDNA was then used as the substrate for PCR reactions using 57 differential display primer combinations. The radio-labelled PCR products were fractionated on a denaturing acrylamide sequencing gels using a Genomyx LR<sup>TM</sup> sequencing apparatus, transferred to 3MM filter paper and dried. The dried acrylamide gels were exposed to autoradiography film (BioMax MR<sup>TM</sup>) overnight. After fractionating the radio-labelled PCR products on denaturing acrylamide gels, it was found that the overwhelming majority of the approximately 18,000 PCR products screened were common to both the wild-type and HD mice (data not shown). One PCR product, amplified using the primers P7 (5'-ATT AAC CCT CAC TAA ATG CTG TAT G-3') and T6 (5'- CAT TAT GCT GAG TGA TAT CTT TTT TTT TCG-3') of approximately 500 bp, was observed in each of three samples derived from the striatum of wild-type mice (FIG. 1). This 500 bp band was absent from the samples derived from the striatum of the HD mice (FIG. 1) and was absent from each of the samples derived from the cortical tissue (data not shown).

FIG. 1 shows the Down-regulated in Huntington's Disease (PDE10A) transcript, identified by differential display RT PCR. A band of approximately 500 bp (arrow) was amplified from cDNA made form 10 week-old wild-type but not 10 week-old HD striatal tissue. Total RNA from individual animals (numbered 1-6) was used as the substrate for the generation of single-stranded cDNA. Animals 1, 2 and 3 were transgenic HD mice. Animals 4, 5 and 6 were wild-type mice.

# EXAMPLE 2 - Cloning of PDE10A

The 500 bp band, designate PDE10Apcr, was excised from the dried gel and rehydrated in 40 μl of H<sub>2</sub>O for 10 min at room temperature. The eluted DNA was subjected to PCR reamplification using the P7 and T6 primers, rTaq polymerase (Pharmacia) and the following conditions: 60" @ 94°C, 19 x (30" @ 94°C, 30" @ 58°C, 120" @ 68°C + 4" per cycle), 7' @ 68°C. The PCR reaction was subjected to agarose gel electrophoresis and the 500 bp band was removed from the gel, extracted from the agarose using the Qiagen gel extraction protocol and cloned into the vector, pGem-T using standard methods. Plasmid DNA was isolated from selected transformants using Qiagen spin columns. The resultant clone was named pPDE10A.

## EXAMPLE 3 - Identification of PDE10A

The cloned insert of pPDE10A was radio-labelled and used as a hybridization probe in northern blot analysis (FIG. 2). Northern blots of total RNA were prepared using the method described in Denovan-Wright et al. (1998). The 500 bp cloned insert of PDE10A was radio-labelled with [α-32P]dCTP (3000 Ci/mmol) using the Ready-to-Go dCTP beads (Pharmacia). Northern blot hybridization, brain tissue preparation and *in situ* hybridization are described in Denovan-Wright et al. (1998). The 500 bp cloned insert of pPDE10A annealed to a transcript of approximately 9.5 kb in total RNA isolated from the striatum of ten week-old wild-type mice.

FIG. 2 demonstrates that PDE10A is expressed in the striatum but not the cortex of wild-type mice and the steady-state levels of PDE10A are reduced in 10 week old transgenic HD mice. The differential expression of PDE10A in HD mice was confirmed by northern blot analysis. The cloned insert of pPDE10A was radio-labelled and used as a hybridization probe in northern blot analysis. The northern blot was prepared by size-fractionating total RNA from the striatum and cortex of three individual 10 week-old HD (1, 2 and 3) and wild-type (4, 5 and 6) mice. Following the hybridization of pPDE10A, the radio-label was removed and the blot was subsequently allowed to hybridize with a probe that detects constituitively expressed cyclophilin. The hybridization pattern of the cyclophilin probe is aligned below the northern blot demonstrating that equivalent amount of RNA were present in each lane. The relative mobility of RNA molecular weight standards (RNA ladder, Gibco BRL) are shown on the left of the northern blot.

The hybridization signal of pPDE10A was significantly lower in the RNA samples derived from the striatum of 10 week-old HD mice. No expression of the PDE10A mRNA was detected in the cortical RNA samples derived from either the wild-type or HD mice.

## EXAMPLE 4 - Sequencing PDE10A

The sequence of the cloned differential display band, pPDE10A, was determined using M13 universal forward and reverse sequencing primers and the T7 sequencing kit (Pharmacia).

The 484 bp cDNA fragment did not have sequence similarity to any Genbank entries.

FIG. 3 shows the nucleotide sequence of the cloned PDE10A differential display product,

pPDE10A. The position of the primers used to amplify the fragment are underlined and labelled. The nucleotide sequence and position of oligonucleotide probes 1 and 2 within the pPDE10A sequence are shown.

## EXAMPLE 5 - Isolation and Characterization of cDNA PDE10A

In order to isolate PDE10A cDNA clones, oligonucleotide probes 1 and 2 were used in 5' and 3' Rapid Amplification of cDNA Ends (RACE) reactions using commercially prepared RACE-ready mouse striatal cDNA (Clontech). Several independent clones were isolated and those that contained the sequence of pPDE10A were selected for further analysis. Each of the 5' RACE clones was identical in sequence over the length that the clones could be aligned. The difference in length between these clones is a result of termination of the original reverse-transcriptase reaction at different positions along the mRNA. No difference in size or sequence was detected between several 3' RACE clones. The longest 5' RACE clone and one 3' RACE clone were completely sequenced using internal primers. The present inventors were able to isolate a very short clone that extended the 5' RACE clone using an internal primer (probe 3, 5'- CTA TTT CAC AAG AGA CTG ACC AGC CAA TAA ATC TC-3'). The compiled sequence of the first PDE10A cDNA clone, named cPDE10A-1 is presented in FIG. 10. cPDE10A-1 is 3235 bp in length. The restriction map of cPDE10A-1 is shown in FIG. 11.

The mRNA that hybridized with pPDE10A was approximately 9.5 kilobases in length. In order to obtain PDE10A cDNA clone that was larger than cPDE10-1, the present inventors screened a mouse brain cDNA library. Several clones were identified that hybridized with

the pPDE10 probe. The sequence of the largest of these cDNA clones, cPDE10-2, was determined. The sequence (FIG. 12) was 5753 base pairs in length. The restriction map of cPDE10-2 is shown in FIG. 13.

cPDE10-1 and cPDE10-2 share sequence identity over 2095 bp. However, the 5' 1142 bp of cPDE10-1 and the 5' 1689 bp of cPDE10-2 are unique to each clone. Clone cPDE10-2 extends 1969 bp in the 3' direction compared to cPDE10-1. A schematic showing the regions of sequence identity and the unique sequences of cPDE10-1 and -2 are shown in FIG. 14.

The compiled sequence of the mouse PDE10 cDNA clone, named cPDE10A, is presented in FIG. 15 with RACEs. A further sequence, without RACEs, is shown in FIG. 19. The coding sequence and restriction map of cPDE10A is shown in FIG. 16, and updated at FIG. 17. FIG. 18 is a restriction map of PDE10A. The coding region has a met initiator commencing at nucleotide 257, with a stop codon ending at nucleotide 2596.

PDE10A was found to have extremely high homology with human PDE10s identified by Loughney et al., WO99/42596, the contents of which are incorporated herein by reference.

## EXAMPLE 6 - Localization of PDE10A in the Brain

In order to identify the coding strand and to localize the transcript in the wild-type mouse brain, two oligonucleotide probes were designed (probe 1, 5'- GAA CAT GTA GCA TAT ACT CCA GAC AAC AGA TCA TAT GG – 3'; probe 2, 5' – CAG CTT CTC CAC AGG AAC ACA GTA ACA AAG AG –3') that were complementary to different regions and

strands of the 484 bp pPDE10A clone. These oligonucleotides were used for *in situ* hybridization analysis. Using high stringency post *in situ* hybridization washes (2 x 30' in 1X SSC @ 58°C, 4 x 15' in 1X SSC @ 58°C, 4 x 15' in 0.5X SSC @ 58°C, 4 x 15' in 0.25X SSC @ 58°C), it was found that oligonucleotide probe 1 annealed with mRNA in the striatum, nucleus accumbens and olfactory tubercule of ten week-old wild-type mice (FIG. 4). The hybridization signal was significantly reduced in the striatum, nucleus accumbens and olfactory tubercle of the 10 week-old HD mice (FIG. 4).

FIG. 5 shows in situ hybridization of probe 1 to coronal (top three sections) and saggital (bottom section) 10 week-old wild-type (WT) and HD mouse brain sections. Specific hybridization of the probe was observed in the striatum, nucleus accumbens and olfactory tubercle of wild-type mice. The top three sections represent the distribution of PDE10A throughout the rostral-caudal axis of the striatum.

The *in situ* hybridization results confirmed the northern blot analysis demonstrating, 1) that the expression of PDE10A mRNA was restricted to the striatum, nucleus accumbens and olfactory tubercle and 2) that the levels of PDE10A mRNA were decreased in HD mice compared to the wild-type. The probe did not anneal with mRNA in any other brain nuclei. No hybridization of oligonucleotide probe 2 was observed in any region of the brain in wild-type or HD mice (Fig. 3). Based on this hybridization, the coding strand, complementary to probe 1, of pPDE10A was defined.

EXAMPLE 7 - Characterization of PDE10

The *in situ* hybridization using oligonucleotide probe 1 demonstrated that PDE10A mRNA levels in the striatum, nucleus accumbens and olfactory tubercule were decreased in ten week- old HD mice. By ten weeks of age, the HD mice all showed motor symptoms including resting tremor and stereotypic involuntary movements. Moreover, these mice immediately clasped their feet together and curled into a tight ball when picked up by their tails.

As the phenotypic signs are progressive over a number of weeks, the present inventors examined whether the PDE10A transcript was ever expressed in the striatum of the HD mice or whether the steady-state levels of the transcript diminished in the striatum in a course that parallelled the development of the motor disorders. Wild-type and HD mice were sacrificed at 5, 7 and 8 weeks of age and their brains were prepared for *in situ* hybridization analysis using probe 1 (FIG. 5).

FIG. 5 shows the levels of PDE10A mRNA decrease in HD mice over the period of time that the HD mice develop abnormal movements and postures. *In situ* hybridization analysis of coronal and saggital sections of wild-type and HD mouse brain using oligonucleotide probe 1 which is complementary to the coding strand of PDE10A. At 5 weeks of age, before the development of motor symptoms, the HD mice express the PDE10A transcript in the same brain nuclei and at the same relative levels as wild-type mice. The steady-state level PDE10A decreases in the striatum, nucleus accumbens and olfactory tubercle from 5 to 10 weeks in the HD but not wild-type mice. By 9 weeks of age, the HD mice have abnormal

movement and posture. The numbers refer to the age in weeks of the wild-type (WT) and Huntington's (HD) transgenic mice.

None of the mice at these ages had overt motor symptoms. Sections taken throughout the rostral-caudal axis of the striatum showed that PDE10A was expressed in the 5 week-old wild-type and HD mice. The relative hybridization of probe 1 did not change in 5, 7, 8 and 10 week-old wild-type mice. The intensity of the hybridization signal appeared to decrease in the striatum, nucleus accumbens and olfactory tubercle of HD mice from 5 to 10 weeks compared to their wild-type litter mates (FIG. 5).

The levels of PDE10A were significantly reduced by 8 weeks of age in the HD mice, using two in situ oligonucleotide probes, one complementary to the 3' UTR, the second complementary to an internal portion of the coding region. The hybridization pattern observed in the wild-type and HD mice was the same for both the probes employed. This analysis demonstrated that there is a reduction in the complete PDE10A mRNA levels during the development of the HD phenotype and not that there was a differential reduction in the PDE10A coding region as compared to the extensive 3' UTR. Moreover, in situ hybridization using the PDE10A-specific probe against neurologically normal and HD human brain tissue demonstrated that there was a decrease in PDE10A levels in human HD patients.

One day old wild-type and HD mice were frozen, sectioned on a cryostat and whole mouse sections were prepared for *in situ* hybridization using probe 1. The same high stringency post-hybridization washing conditions were employed for the one day-old mouse body sections as were used for the adult mouse brain sections. Parallel *in situ* hyridization

experiments using the probe 2 were performed in order to determine the level of non-specific signal in the mouse sections. Probe 1 specifically annealed to the developing striatum (FIG. 6).

FIG. 6 demonstrates that PDE10A is expressed in the developing striatum of one day-old wild-type and HD mice. The sections on the left were subjected to *in situ* hybridization using probe 1. Following hybridization, the sections were counter-stained with cresyl violet to visualize the mouse organs. The signal outside the brain was non-specific as probe 2 and other unrelated control oligonucleotide probes all labelled these tissues.

There was no difference in the pattern of hybridization between the one day-old wild-type and HD mice demonstrating that PDE10A was expressed in the developing brain of both wild-type and HD mice.

Following in situ hybridization, the sections were covered in autoradiographic emulsion, left in the dark to expose for 4 weeks and then developed and viewed under dark-field microscopy or, after counter-staining the sections with cresyl violet to visualize neuronal cell bodies, under bright-field microscopy. Silver grains were observed to be concentrated in the striatum of the wild-type mice. FIG. 7 shows emulsion autoradiography of mouse brain sections following in situ hybridization of probe 1 demonstrated that the PDE10A transcript is expressed in neurons. PDE10A is not homogeneously distributed throughout the mouse striatum. Dark field illumination of the sections after emulsion autoradiography showed that the silver grains were clustered in specific regions of the 10 week old wild-type mouse striatum (A and C). Sections from 10 week old HD mice subjected to identical in situ and

emulsion autoradiographic conditions are shown in B and D. The photomicrographs shown in A and B were viewed using the 10X objective (bar represents  $100 \,\mu m$ ). The micrographs shown in C and D, were viewed under the 20X objective (bar represents  $25 \,\mu m$ ). The insert in panel C is a portion of the section in A and C counter-stained with cresyl violet to visualize the neurons, viewed using the 40X objective under bright filed illumination. Note the distribution of the silver grains over some, but not all, of the striatal neurons as well as being concentrated around clusters of neurons. It appeared that the silver grains were absent from fibre tracks within the striatum. It appeared that PDE10A mRNA was not confined to regions close to the nucleus but was dispersed in cellular processes.

Huntingtin with an expanded polyglutamine tract (htt-HD) is expressed in neurons of the brain and body throughout development and during the lifetime of HD patients (The Huntington's Disease Research Collaborative, 1993; Ross, 1995). Transgenic HD mice express a portion of htt-HD and develop a phenotype with many of the symptoms of HD after a period of normal development and growth (Carter et al., 1999; Cha et al., 1998; Mangiarini et al., 1996). Using differential display RT PCR, northern blot and *in situ* hybridization, we have demonstrated that PDE10A mRNA levels decline in the striatum of HD mice. This specific member of the PDE multigene family is highly expressed in the striatum and olfactory tubercle of mice (Soderling et al., 1999) and rats (Fujishige et al., 1999) and in the caudate and putamen of humans (Fujishige et al., 1999; Loughney et al., 1999). The levels of PDE10A were the same in 5 week old wild-type and HD mice. PDE10A mRNA levels then began to decline and were almost undetectable in the striatum and olfactory tubercle by the time the mice reached 8 weeks of age. This time coincides with the onset of overt motor symptoms in the HD mice indicating that the loss of PDE10A in striatal neurons leads to

dysfunction of the nuclei that control movement. The R6/2 mice develop the HD phenotype in the absence of cell death. The decrease in PDE10A mRNA, therefore, is not due to the loss of PDE10A-expressing cells but rather a change in steady-state RNA levels that occurs due to the expression of mutant huntingtin.

The particular isoform that decreases in HD is PDE10A. PDE10A has been cloned from human lung and fetal brain cDNA libraries (Fujishige et al., 1999; Loughney et al., 1999). It appears that the presence of the expanded polyglutamine tract in huntingtin alters gene expression in the striatum, and that this is the mechanism by which only a small group of neurons in the striatum and cortex are rendered vulnerable to this ubiquitously expressed mutant protein.

EXAMPLE 8 - PDE10A is Highly Conserved Among Mammalian Species

The oligonucleotide (probe 1) complementary to the coding strand of the PDE10A transcript, was also used as an *in situ* hybridization probe against coronal brain sections derived from adult rats. FIG. 8 shows *in situ* hybridization analysis of adult rat brain sections using oligonucleotide probe 1 complementary to the coding-strand of PDE10A revealed that the pattern of expression of PDE10A is the same in rats and mice. The hybridization conditions used to detect the rat homologue of PDE10A in rat brain tissue differed from those used to detect the transcript in mice only in that the stringency of the post-hybridization washes were reduced.

No hybridization was observed in the rat striatum using the post-hybridization washes

employed following the *in situ* hybridization of mouse brain sections. However, when the stringency of the post-hybridization washes was lowered (2 x 60' in 1X SSC @ 42°C, 2 x 60' in 0.5X SSC @ 42°C, 2 x 60' in 0.25X SSC @ room temperature), the PDE10A oligonucleotide probe specifically labelled the adult rat striatum, nucleus accumbens and olfactory tubercule in a pattern indistinguishable from that observed in mouse brain sections. It appears, therefore, that a transcript which shares nucleotide sequence and expression pattern is present in both mice and rats. The evolutionary conservation of PDE10A suggests that it is important for normal function of the basal ganglia.

By northern blot, Fujishige et al. (1999) demonstrated that PDE10A is expressed in human fetal brain. The homology between mouse and human PDE10A is extremely high (data not shown).

## EXAMPLE 9 - Analysis of PDE10A in Genomic DNA

Because the transgenic mice employed in this study have a copy of the human HD 5' UTR, exon 1 with expanded CAG repeat and 262 bp of the intron 1 that has been integrated into an undefined locus of the mouse genome, it was possible that the integration event disrupted the PDE10A gene preventing its expression in the HD mouse striatum. Genomic DNA was isolated from wild-type and HD mice and subjected to Southern blot analysis.

Genomic DNA was isolated from wild-type and HD mice and subjected to Southern blot analysis using pPDE10A as a hybridization probe. The size of the *Bam*HI and *Eco*RI fragments that are present in the transgenic R6/2 line that correspond to the insertion of the

human exon 1 gene fragment are 1.9 and 0.8 (BamHI) and 1.9 (EcoRI) kb. Analysis of the size of the fragments that hybridized with pPDE10A demonstrated that there was no difference in the size of the hybridizing fragments between the wild-type and HD mice. FIG. 9 shows the genomic DNA restriction fragments that hybridized with pPDE10A were the same in wild-type and HD mice. The size of the hybridizing BamHI and EcoRI fragments in each genomic DNA sample is approximately 8 kb and 3 kb, respectively. If the 1.9 kb SacI-EcoRI HD gene fragment integrated into the genome within the BamHI and EcoRI fragments that hybridized with the DHDM cDNA cloned insert, the sizes of the HD hybridizing bands would have been distinct from those of the wild-type. This Southern blot analysis indicates that the gene encoding PDE10A is present as a single-copy in the mouse genome. The numbers at the left of the blot are the relative mobility of molecular weight markers (1 kb ladder, BioRad).

The PDE10A cDNA has since been cloned using a bioinformatics search strategy involving screening of the expressed sequence tag (EST) database for novel PDE cDNA clones. Independently, the mouse PDE10A cDNA was identified after an EST search for novel PDEs with conserved cGMP binding domains (Soderling et al., 1999). The rat isoforms of PDE10A and splice variants have also been described (Fujishige et al., 1999). Human, mouse and rat PDE10A splice variants differ in their 5' untranslated and part of the 5' coding region but are identical in the coding region when the various splice variants are compared within each species. The human, mouse and rat PDE10A coding regions contain 779, 779 and 794 amino acids, respectively, encoding a protein of approximately 88.5 kDa.

EXAMPLE 10 - Distribution of PDE10A

In mouse, PDE10A mRNA was detected in testis and to a much lesser extent in brain but not in heart, spleen, lung, liver, skeletal muscle, kidney, ovary, pancreas, smooth muscle, eye or in total RNA isolated from 7, 11, 15 or 17 day old embryo (Soderling et al., 1999). This data agrees with the PDE10A mRNA pattern of distribution that we observed in wild-type and pre-symptomatic HD mice. In mice, two different size transcripts are detected in northern blots using the coding region as a probe. In mouse testis, the most abundant transcript is approximately 4 kb. A 9.5 kb transcript was also detected in mouse testis. It appears that the most abundant transcript in mouse brain is 9.5 k. Similarly, two sized PDE10A transcripts were observed in rats, however, it appears that, in rat, the 4 kb mRNA is expressed exclusively in testis while the 9.5 kb mRNA is expressed exclusively in brain (Fujishige et al., 1999). Within the brain, the rat PDE10A mRNA was expressed in striatum and olfactory tubercle and not cortex, cerebellum, hippocampus, midbrain or brainstem. In humans, PDE10A is expressed in the caudate, putamen and testis. As was observed in rodents. mRNAs of approximately 4 and 10 kb hybridized with the PDE10A probe. Again, it appears that, although both sized transcripts are present in brain and testis, the larger mRNA is predominant in the caudate and putamen and the smaller sized transcript is present in the testis. Each of the mouse, rat and human PDE10A sequences are not longer than 4 kb and span the coding region and parts of the 3' UTR. The difference in abundance of the short and long transcript in the testis and brain, respectively, in all three species suggest that the 3' UTR functions to provide transcript stability in the brain. As such, we present the complete sequence of the brain-specific transcript of PDE10A derived from mouse.

EXAMPLE 11 - Modulating Activity of PDE10A Using cGMP-PDE Activity

Cyclic nucleotides are the predominant second messengers that activate cellular signaling pathways (Beavo, 1995; Conti and Jin, 1999). The concentration of intracellular cyclic nucleotides is dependent on their rate of synthesis by adenyl and guanyl synthase, the rate of efflux from the cell, and the rate of degradation. PDEs hydrolyze cAMP and cGMP limiting both the duration and amplitude of the cyclic nucleotide signal (Beavo, 1995; Conti and Jin, 1999). In mammals, PDEs are encoded by a large multigene family. The various PDE family members have tissue-specific patterns of expression (Conti and Jin, 1999). PDEs have also been described in Caenorhabditis, Drosophila, Dictyostelium, Saccharomyces, Candida and Vibrio species demonstrating that this enzyme has been conserved throughout evolution. In mammals, PDEs are encoded by at least 10 gene families, each composed of one or more genes. In addition, numerous splice variants of individual gene family members have been described. These splice variants alter the 5' domain of the protein but share identical nucleotide binding and catalytic domains. The catalytic domain, found in the carboxyterminus of the enzyme, is ~ 275 amino acids and highly conserved in amino acid sequence in all PDEs. In total, it appears that there are ~50 PDEs expressed within the mammalian body. Some PDEs are expressed in multiple tissues while others have a very limited tissue-specific distribution (Conti and Jin, 1999).

PDE gene families differ with respect to their affinity for cAMP and cGMP and their dependence on calcium and calmodulin (Beavo, 1995). Moreover, some PDEs are inhibited or activated by binding cyclic nucleotides to a non-hydrolytic site. For example, PDE2A has a lower  $K_m$  for cGMP than cAMP although it hydrolysed both nucleotides. The binding of

cGMP to an allosteric activator site within PDE2 enhances the rate of catalysis of cAMP. PDE2 is, therefore, a cGMP-stimulated cGMP and cAMP phosphodiesterase (Beavo, 1995). Conversely, the affinity of PDE4 for cAMP is much greater than for cGMP and PDE4 activity is not affected by cGMP or calmodulin (Beavo, 1995). The differences in substrate preference, modulation of activity and tissue-specific patterns of expression suggest that subtle alterations in the relative levels of cAMP and cGMP mediated through the action of various PDEs lead to a wide range of responses to extracellular signals.

cGMP-PDE activity of compounds is measured using a one-step assay adapted from Wells at al. (Wells, J. N., Baird, C. E., Wu, Y. J. and Hardman, J. G., *Biochim. Biophys. Acta* 384:430 (1975)) and adopted by Beavo et al, U.S. Patent No. 6,037,119. The reaction medium contains 50 mM Tris-HCl, pH 7.5, 5 mM Mg-acetate, 250 ug/ml 5'-Nucleotidase, 1 mM EGTA and 0.15 uM 8-[H³]-cGMP. The enzyme used is a human recombinant PDE V (ICOS, Seattle U.S.A.).

Compounds of interest are dissolved in DMSO finally present at 2% in the assay. The incubation time was 30 minutes during which the total substrate conversion did not exceed 30%.

The IC <sub>50</sub> values for the compounds examined are determined from concentration-response curves using typically concentrations ranging from 10 nM to 10 uM. Tests against other PDE enzymes using standard methodology also show compounds highly selective for the cGMP specific PDE enzyme.

Rat aortic smooth muscle cells (RSMC) are prepared according to Chamley et al. in *Cell Tissue Res.* 177:503-522 (1977) and used between the 10th and 25th passage at confluence in 24-well culture dishes. Culture media is aspirated and replaced with PBS (0.5 ml) containing the compound tested at the appropriate concentration. After 30 minutes at 37° C, particulates guanylate cyclase are stimulated by addition of ANF (100 nM) for 10 minutes. At the end of incubation, the medium is withdrawn and two extractions were performed by addition of 65% ethanol (0.25 ml). The two ethanolic extracts are pooled and evaporated until dryness, using a Speed-vat system. c-GMP was measured after acetylation by scintillation proximity immunoassay (AMERSHAM). The EC<sub>50</sub> values are expressed as the dose giving half of the stimulation at saturating concentrations.

## EXAMPLE 12 - Selected Modulators of PDE10A Activity

The catalytic domain of PDE10A is most similar in amino acid sequence to PDE5A, PDE2A, PDE6B and PDE6A. These members of the PDE family each contain a cGMP binding sequence that is not observed in other PDE family members. The non-catalytic cGMP binding sites (GAF) domains found in PDE2, 5 and 6 are also found in PDE10. At least for PDE2, this site acts as an allosteric activator for cAMP hydrolytic activity. The GAF domain of PDE10A binds other small molecules that act as allosteric activators. PDE10A is a cAMP and cAMP-inhibited cGMP PDE (Fujishige et al., 1999; Fujishige et al., 1999; Loughney et al., 1999; Soderling et al., 1999).

Attenuation of the production of cAMP, may ameliorate the symptoms of HD and positively affect gene expression. Pharmaceutically acceptable modulators of cAMP include quinpirole,

alloxan, miconazole nitrate, MDL-12330A, and tetracyline derivatives such as demeclocycline and minocycline.

Compounds which are potent and selective modulators of cGMP-specific PDE, and are useful in a variety of therapeutic areas are taught by Daugan et al, U.S. patent No. 5,981,527, PCT publication No. WO 00/15639 to Icos Corporation and PCT publication No. WO 00/15228 to Icos Corporation, which are incorporated herein by reference. Such compounds include, for example:

(6R,12aR)-2,3,6,7,12,12a-Hexahydro-6-(5-benzofuranyl)-2-methyl-pyrazino[2', 1':6,1]pyrido[3,4-b]indole-1,4-dione,

(6R,12aR)-2,3,6,7,12,12a-Hexahydro-6-(5-benzofuranyl)-pyrazino[2',1':6,1]py rido[3,4-]indole-1,4-dione,

(6R,12aR)-2,3,6,7,12,12a-Hexahydro-6-(5-benzofuranyl)-2-isopropyl-pyrazino[2',1':6,1]pyrido[3,4-b]indole-1,4-dione,

(3S,6R,12aR)-2,3,6,7,12,12a-Hexahydro-6-(5-benzofuranyl)-3-methyl-pyrazino[2',1':6,1]pyrido[3,4-b]indole-1,4-dione, and

(3S,6R,12aR)-2,3,6,7,12,12a-Hexahydro-6-(5-benzofuranyl)-2,3-dimethyl-pyraz ino[2',1':6,1]pyrido[3,4-b]indole-1,4-dione.

PDE1B1 is expressed throughout the brain and is most abundant in the striatum, nucleus accumbens and olfactory tubercle (Polli and Kincaid, 1994; Yan et al., 1994). PDE1B is a cGMP, Ca/calmodulin-dependent PDE. Therefore, PDE1B and 10A are both expressed in the majority, but not all, striatal neurons and, it is likely that both genes are co-expressed in a subset of striatal projection neurons. Selective inhibitors for PDE1 include KS-505, IC224,

and SCH 51866. Of these inhibitors, it appears that SCH 51866 has a ten-fold higher Km for PDE1 than PDE10 (Soderling et al., 1999). The non-specific PDE inhibitor IBMX is a potent inhibitor of PDE10A. Dipyridamole and SCH51866 had the highest potency of inhibitors tested on PDE10A activity. Dipyridamole was considered to be a PDE5- and PDE6-specific inhibitor, however, the Km for dipyridamole is 10 times higher for PDE10A than the other PDEs (Soderling et al., 1999). Selective inhibitors of PDE5, 2, 3 and 4 had much greater IC50 for PDE10 (Soderling et al., 1999).

## EXAMPLE 13 - Clinical use of PDE10A Modulator

A 38 year-old female was admitted to hospital from a long-term care facility due to progressive deterioration of her physical and mental symptoms caused by Huntington's disease. The patient had been diagnosed with Huntington's disease at age 26. Prior to admission to the hospital, she had become increasingly aggressive and uncooperative.

Moreover, there appeared to be an increase in the number of psychotic episodes. SPECT showed no abnormality of brain blood flow but MRI showed bilateral caudate atrophy as well as global atrophy of the cerebrum and corpus callosum.

The patient had been stable for a number of years on the antipsycotic haloperidol (3 mg/day). For the last two years, the haloperidol had been replaced by olanzapine (2.5-7.5 mg/day).

Minocycline, a tetracycline derivative, was administered at 50 mg twice daily for 7 days, followed by 100 mg twice daily for 7 days and finally 200 mg twice daily for 5 weeks. After 5 weeks of 200 mg twice daily minocycline administration, there was a mild improvement

compared to the baseline clinical global assessment made at the time of admission. The minocycline treatment was suspended for 7 days. Due to a significant increase in the number of aggressive incidence and decrease in cooperativity, minocycline (200 mg twice daily) treatment was resumed. The patient responded within 3 days to the resumed minocycline-treatment with a return to mild-improvement compared to the baseline clinical global assessment made at the time of admission. Minocycline (200 mg twice daily) treatment will continue indefinitely. The improvement in behaviour and decrease in apparent psychosis has allowed for the transfer of the patient from the acute care facility back to long-term care.

While the present invention has been described in terms of specific embodiments, it is understood that variations and modifications will occur to those skilled in the art.

Accordingly, only such limitations as appear in the appended claims should be placed on the invention.

Non-patent Literature Cited

- Atschul, S. F. et al. (1990), J. Molec. Biol. 215: 403.
- Babity, J.M., Armstrong, J.N., Plumier, J.-C., Currie, R.W., and Robertson, H.A. (1997a) A novel seizure-induced synaptotagmin gene identified by differential display. *Proc. Nat. Acad. Sci. U.S.A.* 94: 2638-2641.
- Babity, J.M., Newton, R.N., Guido, M.E. and Robertson, H.A. (1997b) The Application Of Differential Display To The Brain: Adaptations For The Study Of Heterogeneous Tissue.

  Methods Mol Biol, 85 (1997) 285-95.
- Barney (1991) in PCR Methods and Applications, Aug., Vol. 1(1), page 5.
- Beavo, J. (1995). Cyclic nucleotide phosphodiesterase: functional implications of multiple isoforms. Physiol. Rev. 75, 725-748
- Beretta, S., Robertson, H. and Graybiel, A. (1992) Dopamine and glutamate agonists stimulate neuron-specific expression of Fos-like protein in the striatum. *J. Neurophysiol.* 68, 767-777
- Bennett, D. et al., (1995) Journal of Molecular Recognition, Vol. 8 52-58.
- Berke, J.D., Paletzki, R.F., Aronson, G.J., Hyman, S.E. and Gerfen, C.R. (1998) A Complex Program of Striatal Gene Expression Induced by Dopaminergic Stimulation. *J. Neurosci.* 18: 5301-5310.
- Caine E.D., Hunt R.D., Weingartner H., Ebert M.H. (1978) Huntington's dementia.

  Clinical and neuropsychological features. *Arch-Gen-Psychiatry*. 35, 377-84.
- Carillo, H., and Lipman, D.(1988), SIAM J. Applied Math., 48: 1073.
- Carter, R. J., Lione, L. A., Humby, T., Mangiarini, L., Mahal, A., Bates, G. P., Dunnet, S. B., and Morton, A. J. (Apr. 1999). Characterization of progressive motor deficits in mice transgenic for the human Huntington's diease mutation. J Neurosci 19, 3248-57

Cha, J.-H. J., Kosinski, C.M., Kerner, J.A., Alsdorf, S.A., Mangiarini, L., Davies, S.W., Penny, John B., Bates, G.P., Young, A.B. (1998) Altered brain neurotransmitter receptors in transgenic mice expressing a portion of an abnormal human Huntington disease gene. *Proc. Natl. Acad. Sci. USA* 95: 6480-6485.

- Conolly, B. A. (1987) Nucleic Acids Res. 15:15(7): 3131.
- Conti, M., and Jin, S. L. (1999). The molecular biology of cyclic nucleotide phosphodiesterases. Prog Nucleic Acid Res Mol Biol 63, 1-38
- Crino, P., Khodakhah, K., Becker, K., Ginsberg, S., Hemby, S. and Eberwine, J. (1998)

  Presence and phosphorylation of transcription factors in developing dendrites. *Proc.*Natl. Acad. Sci. USA 95: 2313-2318
- Davies, S.W., Turmaine, M., Cozens, B.A., DiFiglia, M., Sharp, A.H., Ross, C.A., Scherzinger, E., Wanker, E.E., Mangiarini, L. and Bates, G.P. (1997) Formation of neuronal intranuclear inclusions underlies the neurological dysfunction in mice transgenic for the HD mutation. *Cell* 90, 537-548.
- Davis et al., (1986) Basic Methods in Molecular Biology.
- Denovan-Wright, E.M., Newton, R.A., Armstrong, J.N., Babity, J.M., Robertson, H.A. (1998) Acute administration of cocaine, but not amphetamine, increases the level of synaptotagmin IV mRNA in the dorsal striatum of rat. *Mol. Brain Res.* 55, 350-354.
- Denovan-Wright, E.M., Gilby, K.L., Howlett, S.E. and Robertson, H.A. (1999) Cloning of differentially expressed brain cDNA. In *Differential Display A Practical Approach*, H.A. Robertson and R.A. Leslie, Eds. Oxford University Press (in press).
- Devereux, J., et al.(1984), Nucleic Acids Research 12(1): 387.
- Douglass, J., McKinzie, A.A. and Couceyro, P. (1995) PCR differential display identifies a rat brain mRNA that is transcriptionally regulated by cocaine and amphetamine. J.

- Neurosci. 15, 2471-2481.
- Douglas, J. and Daoud, S. (1996) Characterization of the human cDNA and genomic DNA encoding CART: a cocaine- and amphetamine-regulated transcript. *Gene* 169, 241-245
- Eberwine, J. (1996) Amplification of mRNA populations using aRNA generated from immobilized oligo(dt)-T7 primed cDNA. *Biotechniques* 20: 584-589
- Eberwine, J., Yeh, H., Miyashiro, K., Cao, Y., Nair, S., Finnell, R., Zettel, M., Coleman, P., (1992) Analysis of gene expression in single live neurons. *Proc. Natl. Acad. Sci. USA* 89: 3010-3014
- Fujishige, K., Kotera, J., Michibata, H., Yuasa, K., Takebayashi, S., Okumura, K., and Omori, K. (June 25, 1999). Cloning and characterization of a novel human phosphodiesterase that hydrolyzes both cAMP and cGMP (PDE10A). J Biol Chem 274, 18438-45
- Glass, M., Faull, R.L.M. and Dragunow, M. (1993) Loss of cannabinoid receptors in the substantia nigra in Huntington's disease. *Neuroscience* 56: 523-527 Good et al (1977), *Nucl. Acid Res* 4:2157.
- Gribskov, M. and Devereux, J., eds., Sequence Analysis Primer, M Stockton Press, New York, 1991.
- Griffin, A. M., and Griffin, H. G., eds., Computer Analysis of Sequence Data, Part I, Humana Press, New Jersey, 1994.
- Group THDCR (1993) A novel gene containing a trinucleotide repeat that is extended and unstable on Huntington's disease chromosomes. *Cell* 72: 971-983.
- Hardy, J. and Gwinn-Hardy, K. (1998) Genetic classification of primary neurodegeneraive disease. *Science* 282: 1075-1079
- Hinnen et al., (1978) PNAS USA 75:1929-1933.

Houbenweyl, (1987) Methods of Organic Chemistry, ed. E. Wansch, Vol. 15 I and II, Thieme, Stuttgart.

- Innis, M. A. and D. H. Gelfand (1989), PCR Protocols, A guide to Methods and Applications, M. A. Innis, D. H. Gelfand, J. J. Shinsky and T. J. White eds, pp 3-12, Academic Press.
- Johanson, K. et al.(1995), The Journal of Biological Chemistry, Vol. 270, No. 16, pp 9459-9471.
- Klement, I.A., Skinner, P.J., Kaytor, M.D., Yi, H., Hersch, S.M., Clark, H.B., Zoghbi, H.Y. and Orr, H.T. (1998) Ataxin-1 nuclear localization and aggregation: role in polyglutamine-induced disease in *SCA1* transgenic mice. *Cell* 95: 41-53.
- Lesk, A. M., ed., *Computational Molecular Biology*, Oxford University Press, New York, 1988.
- Liang, P. and Pardee, A.B. (1992) Differential display of eukaryotic messenger RNA by means of the polymerase chain reaction. *Science* 257: 967-971.
- Livesey, F.J., O'Brien, J.A., Li, M., Smith, A.G., Murphy, L.J. and Hunt, S.P. (1997) A

  Schwann cell mitogen accompanying regeneration of motor neurons. *Nature* 390: 614-618.
- Livesey, F.J. and Hunt, S.P. (1996) Identifying changes in gene expression in the nervous system: mRNA differential display. *Trends Neurosci.* 19: 84-88
- Loughney, K., Snyder, P. B., Uher, L., Rosman, G. J., Ferguson, K., and Florio, V. A. (June 1999). Isolation and characterization of PDE10A, a novel human 3', 5'-cyclic nucleotide phosphodiesterase. Gene 234, 109-17
- Ludlow, C.L., Connor, N.P., Bassich, C.J. (1987) Speech timing in Parkinson's and Huntington's disease. *Brain-Lang*, 32, 195-214.
- Mangiarini, L., Sathasivam, K., Seller, M., Cozens, B., Harper, A., Hetherington, C. Lawton,

M., Trottier, Y., Lehrach, H., Davies, S. W. and Bates, G. P. (1996) Exon 1 of the HD gene with an expanded CAG repeat is sufficient to cause a progressive neurological phenotype in transgenic mice. *Cell* 87: 493-506.

- Merrifield, (1964), J. Am. Chem. Assoc. 85:2149-2154.
- Okano, J., (1988) Neurochem. 56: 560.
- Pardee, A.B. (1997) Complete genome expression monitoring: the human race. *Nat. Biotechnol.* 15: 1343-1344
- Polli, J.W., and Kincaid, R. L. (1994). Expression of a calmodulin-dependent phosphodiesterase isoform (PDE1B1) correlates with brain regions having extensive dopaminergic innervation. J. Neurosci. 14, 1251-1261
- Qu,-D., Ludwig,-D.S., Gammeltoft, S., Piper, M., Pelleymounter, M.A., Cullen, M.J. Mathes, W.F., Przypek, R., Kanarek, R. and Maratos-Flier, E. (1996) A role for melanin-concentrating hormone in the central regulation of feeding behaviour. *Nature* 380: 243-247.
- Richfield, E.K. and Herkenham, M. (1994) Selective vulnerability in Huntington's disease: preferenctial loss of cannabinoid receptors in lateral globus pallidus. *Ann. Neurol.* 36: 577-584.
- Ross, C.A., (1997) Intranuclear neuronal inclusions: a common pathogenic mechanism for glutamine-repeat neurodegenerative diseases? *Neuron* 19: 1147-1150.
- Saiki et al., (1986) Nature, 324: 163-166.
- Sambrook, et al, (1989), *Molecular Cloning, A Laboratory Manual*, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y.
- Saudou, F., Finkbeiner, S., Devys, D. and Greenberg, M.E. (1998) Huntingtin acts in the nucleus to induce apoptosis but death does not correlate with the formation of intranuclear

inclusions. Cell 95: 55-66.

- Sisodia, S.S., (1998) Nuclear inclusions in glutamine repeat disorders: are they pernicious, coincidental or beneficial? *Cell* 95: 1-4.
- Smith, D. W., ed., Biocomputing: Informatics and Genome Projects, Academic Press, New York, 1993.
- Soderling, S. H., Bayuga, S. J., and Beavo. J. A. (June 1999). Isolation and characterization of a dual-substrate phosphodiesterase gene family: PDE10A. Proc Natl Acad Sci U S A 96, 7071-6
- von Heinje, G., Sequence Analysis in Molecular Biology, Academic Press, 1987.
- Yan, C., Bentley, J. K. Sonnenburg, W. K., and Beavo, J. A. (1994). Differential expression of the 61 kDa a 63 kDa calmodulin-dependent phosphodiesterase in the mouse brain. J. Neurosci. 14, 973-984
- Young, A.B., Penney, J.B., Starosta-Rubinstein, S., Markel, D.S., Berent, S., Giordani,
  B., Ehrenkaufer, R., Jewett, D., Hichwa, R. (1986) PET scan investigations of
  Huntington's disease: cerebral metabolic correlates of neurological features and
  functional decline. *Ann-Neurol*. 20, 296-303.

#### We claim:

1. A composition for treating a CAG repeat disorder comprising a compound which modulates PDE10A expression and a pharmaceutically acceptable carrier.

- 2. A composition as claimed in claim 1, wherein said compound is selected from the group consisting of: quinpirole, alloxan, miconazole nitrate MDL-12330A, and tetracyline derivatives such as demeclocycline.
- 3. A composition as claimed in claim 1, wherein said compound is selected from the group consisting of:

(6R,12aR)-2,3,6,7,12,12a-Hexahydro-6-(5-benzofuranyl)-2-methyl-pyrazino[2', 1':6,1]pyrido[3,4-b]indole-1,4-dione,

(6R,12aR)-2,3,6,7,12,12a-Hexahydro-6-(5-benzofuranyl)-pyrazino[2',1':6,1]py rido[3,4-lindole-1,4-dione,

(6R,12aR)-2,3,6,7,12,12a-Hexahydro-6-(5-benzofuranyl)-2-isopropyl-pyrazino[2',1':6,1]pyrido[3,4-b]indole-1,4-dione,

(3S,6R,12aR)-2,3,6,7,12,12a-Hexahydro-6-(5-benzofuranyl)-3-methyl-pyrazino[2',1':6,1]pyrido[3,4-b]indole-1,4-dione,

(3S,6R,12aR)-2,3,6,7,12,12a-Hexahydro-6-(5-benzofuranyl)-2,3-dimethyl-pyraz ino[2',1':6,1]pyrido[3,4-b]indole-1,4-dione.

4. A composition as claimed in claim 1, wherein said compound is selected from the group consisting of: KS-505, IC224,SCH 51866, IBMX and Dipyridamole.

5. A composition as claimed in any one of claims 1 to 4, wherein said disorder is Huntington's disease.

- 6. The use of a composition as claimed in any one of claims 1 to 5 for treating a CAG repeat disorder comprising administering said composition to a subject in need of such treatment.
- 7. The use of a composition of claim 6 for treating Huntington's disease comprising administering said composition to a subject in need of such treatment.
- 8. A method for identifying a compound which inhibits or promotes a CAG repeat disorder, comprising the steps of:
- (a) selecting a control animal having PDE10A and a test animal having PDE10A;
- (b) treating said test animal using a compound; and,
- (c) determining the relative quantity of RNA corresponding to PDE10A, as between said animals.
- 9. A method of claim 8, wherein said animal is a mammal.
- 10. A method of claim 9, wherein said mammal is a mouse.
- 11. A method of claim 10, wherein said mouse is R6/2 transgenic mouse.
- 12. A method of any one of claims 8 to 11, wherein said CAG repeat disorder is Huntington's disease.

13. A method for identifying a compound which inhibits or promotes a CAG repeat disorder, comprising the steps of:

- (a) selecting a host cell containing PDE10A;
- (b) cloning said host cell and separating said clones into a test group and a control group;
- (c) treating said test group using a compound; and
- (c) determining the relative quantity of RNA corresponding to PDE10A, as between said test group and said control group.
- 14. A method of claim 13, wherein said CAG repeat disorder is Huntington's disease.
- 15. A method for detecting the presence of or the predisposition for a CAG repeat disorder, said method comprising determining the level of expression of RNA corresponding to PDE10A in an individual relative to a predetermined control level of expression, wherein a decreased expression of said RNA as compared to said control is indicative of a CAG repeat disorder.
- 16. A method of claim 15, wherein said CAG repeat disorder is Huntington's disease.
- 17. A method of claim 15 or 16, wherein said expression is measured by in situ hybridization.
- 18. A method of claim 15 or 16, wherein said expression is measured using a polymerase chain reaction.

19. A method of claim 15 or 16, wherein said expression is measured using a DNA fingerprinting technique.

HD WT

1 2 3 4 5 6



Figure 1

Striatum Cortex

HD WT HD WT

123456123456

9.5

7.5



4.4

2.4

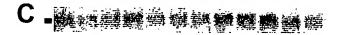
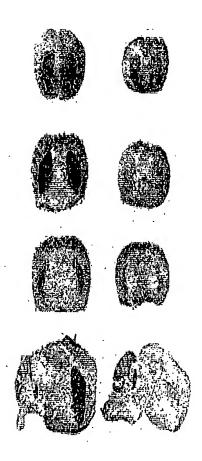


Figure 2

## Figure 3

5 1		11		21		31	•	41	
,	TGTATGGGA	ATAG	TGTTTC	CATAT	GATCT	GTTG:	CTGGA	TAT	ATGCTAC
,	ACATACCCT	TATC	ACAAAE	<b>STATA</b>	CTAGA	CAAC	AGACCTO	ATA	TACGATG
									probe 1
5 '		61		71		81		91	
51	ATGTTCATT TACAAGTAA	TACT	GTACAA	AAACC	CAGTG(	CAGCI	GATGAT	GCA	AAGCAGT
	IACAAGIAA	AIGA	CAIGII.	1100	GICAC	3 I CGF	CIACIA	CGI	IICGICA
5 '		11		21		31	•	41	
101	CTCTCTCTGT	rgta	CAGTGC	CCAC	CTATT	<b>LAAA7</b>	ATCACO	TAC	TTGCCCA
101	GAGAGAGACA	ACAT	GTCACGO	GGTG	GATAA	ATTTI	TAGTGC	ATG	AACGGGT
5 '		61		71		81		91	
151	GAACACTGTC CTTGTGACAC	AAAE	CACTTA	CATA	AGAAC	AAACG	CAGCGI	'CTG	GATTCTT
TOT	CTTGTGACAC	CTTT	GTGAATI	GTAT	rcttgi	rTTGC	GTCGCA	GAC	CTAAGAA
5 '	TCCAAGGAGA	11	probe 2	21		31		41	
201	TCCAAGGAGA	AGEV	GCTTTCT	CCAC	AGGAAC	CACAG	TAACAA	AAG	AGGTCCG
201	AGGTTCCTCT	CGT	CGAAAGA	GGTG	CCTTG	FIGTO	ATTGTT	TTC	ICCAGGC
·	•								
5 '		61	· · · · · · · · · · · · · · · · · · ·	71		81		91	
251	CCGCCATCCA	CAC	CCAGCCA	AGAC	ACCTCA	GAGG	CCATAG	GGA	CAACCTC
	GGCGGTAGGT	GIGG	3GTCGGT	'TCTGT	rggag'i	CICC	GGTATC	CCL	JTTGGAG
5 '		11		21		31		41	
<b>D</b> .	ammaamaaaa		aamaam		א מממר		CCTCCC		1 7 CTC 1 TT
301	CTTGCTGGCC		ACCIGCI	COTO	᠉ᡎᢗ᠊ᢗᢗᢗ	CACA		THO CO	PTCD CTD
	GAACGACCGG	1141	GGACGA	CCIC	310000	.GIGI	CCAGGG	Teg.	IIGACIA
5 '		61		71.		81		91	
	COTONOTCON		יייכייככא	מרכא א	АСССТ		מממרדר		יייית אמ
351	CCTCAGTGGA GGAGTCACCT	7000	'ACACCT	CCCTT	MUGGE MUGGE	מידים	CCCGAG	AGAZ	יייייטעעעע
	GGAG1 CACC1	ACCC	HOHOGI	COO11	. 1 00021		00000		
5 '		1.1		21		31		41	
	GGGAAAGAAA	GAAT	TTCAAG	CTTAT	GATAT	CCAA	TATTAT	TAT	AGTTGAT
01	GGGAAAGAAA CCCTTTCTTT	CTTA	AAGTTC	GAATA	CTATA	GGTT.	ATAATA	ATA	CAACTA
5 '		61		71		81		91	
<b>-</b> -	GAGTTAGTAA. CTCAATCATT	ATTC	CAAAAA.	AAAAA					
ЭŢ	CTCAATCATT'	TAAG	CTTTTTT	$\mathbf{T}\mathbf{T}\mathbf{T}\mathbf{T}\mathbf{T}\mathbf{T}$	ı				



WT HD

Figure 4

WO 01/24781

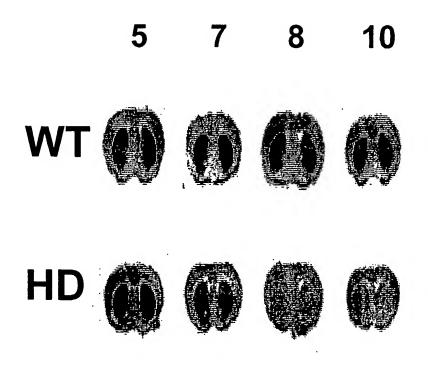


Figure 5

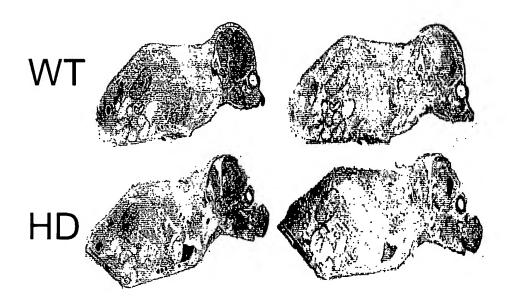


Figure 6

WO 01/24781

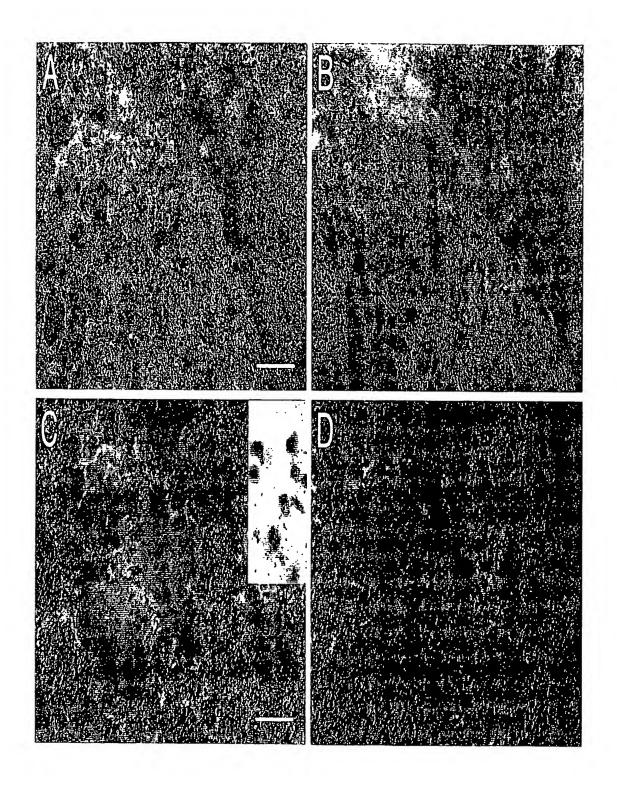


Figure 7



Figure 8

### BamHI EcoRI

### W1W2 H1H2 W1W2 H1H2

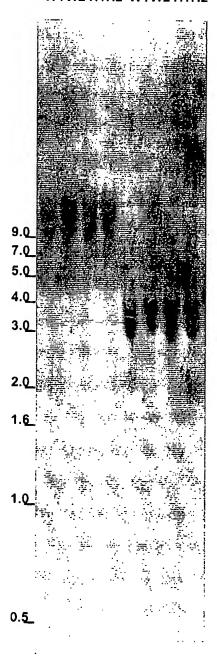


Figure 9

## Figure 10

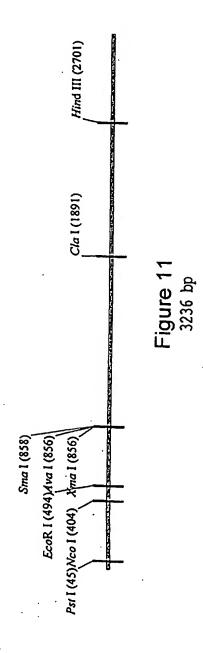
5			31	41	
	CACTGAAGCTGGTCCA GTGACTTCGACCAGGT	CGTCTATAA GCAGATATT	ACAGGTGACAC TGTCCACTGTG	IGGCTGCAGCA ACCGACGTCGT	AAA. TTT
5 '	61	71	81	91	
5	1 AGCCATTCGATCCACA TCGGTAAGCTAGGTGT	CAAA TTGAT GTTT AACTA	CTTCTATCATC' GAAGATAGTAG	TTGG AATCTGA AACC TTAGACT	RTA AAT
5'	11	21	31	41	
10:	GCAGGGAGGAGCAGTA CGȚCCCTCCTCGTCAT	TGTA AGACGI ACAT TCTGC	ACCGTTTAATT( CGGCAAATTAA(	CAGG CATTCCG STCCGTAAGGC	AAG TTC
· 5 ·	61		81	91	
151	GCATGAGCGCATGGAT CGTACTCGCGTACCTA				TTA AAT
5'	11	21	31	. 41	
201	GGGAAACCTATGACGG CCCTTTGGATACTGCC	ACTG TTTTTC CGACAAAAAC	CTGT AGAAGTA GACA TCTTCAT	GGGATTTTAC CCCTAAAATG	AGA TCT
5 '	61	71	81	91	
251	AGTCTCCTTGAATTTGC TCAGAGGAACTTAAACC	CCTGCCTGG GGACGGACC	GGCAGTTTTGC CCGTCAAAACG	'AGA GGAACCT( 'TCT CCTTGGA(	GCC CGG
5 '	11	21	31	41	
301	AGAGATTTATTGGCTGG TCTCTAAATAACCGACG	STCAGTCTCT LAGT CAGAGA	TGTGAAATAGT ACACTTTATCA	ATCATGTGAGI TAGTACACTC	AAA TTT
5 '	61	71	81	91	
351	CAGTTTGTAGAAAAAAA GTCAAACATCTTTTTT	CTATACCTG GATATGGAC	GGAA GACCTTT CCTT CTGGAAA	GCAACATTGT1 CGTTGTAACAA	AGG AGC
5 '	11	21	31	41	
401	TTCCATGGGCCAAGACT AAGGTACCCGGTTCTGA	CAGTTAGGA GTCAATCCT	GGCA TAAATCT CCGT ATTTAGA	GCCCGGAATAA CGGGCCTTATT	AAC
5'	61	71	81	91	
451	TAGGCCAGGATACAGCC ATCCGGTCCTATGTCGG	ATGTTTAGT TACAAATCA	PAATAATTTGG ATTA TTAAACC	PTT TAGAATTO AAAATCTTAAG	AC
5'	11	21	31	41	
501	ACAGGCAGGATTGGTTT TGTCCGTCCTAACCAAA	TTTTGTGTCT AAAACACAGA	TTGG CAAGTGG! ACCGTTCACC!	AGCATATTTAA PCGTATAAATT	CA GT
·5 <sup>1</sup>	· 61	71	81 .	91	
551	TACAGGCATGGGAATCCT ATGTCCGTACCCTTAGGA	rgcctcttac Acggagaatc	CTTTTCCCAC( CAAAAGGGTG(	CTCTTGTCTC EGAGAACAGAG	AC
5 '	· 11	21	31	41	
601	CAAGTTTTTTCTCTCCAA	AGGTTTCCA	GGAATTTCTCF	TTAATGGCTG	AT

5 '	61		81	91	
651 GCAA	CTTAGTGAATA	ATAA TGAATA	TAAA CAATGC	TCACCTCACCAA AGTGGAGTGGTT	AA
OD CGTTI	GAATCACTTAT	TATTACTTAT	ratificitacs.	AGTGGAGTGGTT	.T.T.
5 '	11	21	31	41	
TTATA	TTATTTGCAGT	CATT TGTGAT	TAACA CAAATT	TTAT CGCAATGG AATA GCGTTACC	TT
701 AATAT	AATAA ACGTCA	GTAAACACTA	TTGTGTTTAA	AATAGCGTTACC	AA
	61	71	81	91	
5 '	61 ייייא א יייייער מייינר	7.1 ንፐጋሬጋሬ ጋሬጋን	LTCTATTSSTE	TTGTTGTGGTTG	тт
751 TAATA	AATTAAACACC	GTGTGTGAC	ACCAATAGAA	TTGTTGTGGTTG AACAACACCAAC	ĀĀ
5 <b>'</b>	11	. 21	31	41	B 677
801 TCTGA	GAAAA TGTTCT.	IGGATATGTA	AGTGCCAATA( TCDCCCTTDT(	CCAGTGTGAAGT. EGTCACACTTCA	AΤ
AGACT	CTTTTACAAGAA	ACCIMIACAI	ICACGGIIAI	2010110101101	
5'	61	71	81	. <b>91</b>	
TGATO	CCGGG CAGCAA	ATA CAGCCT	AAGGTTTGTA	AACATCAATTCT. FTGTAGTTAAGA	AΤ
851 ACTAG	GGCCCGTCGTTT	TATGTCGGA	TTCCAAACAT	ITGTAGTTAAGA'	ΓA
g i	. 11	21	31	41	
5 '	ተተ ተተ	CCTGAGAAG	CTGCGGGGCA	TGTAAAGTAAA	GT
901 GAGTC	AAGTAGTCTCCC	CGGACTCTTC	GACGCCCCGT	STGTAAAGTAAA CACATTTCATTT	CA
	•				
5'	61 ~~~~~	71	81	91 AGTGACCAGTGC	ፐር፤
951 TACCA	CCCCA CCACCAC	CAGTCAGCCIC	CCGCCTGRAGE GGCGGACTTCT	CACTGGTCACG	ĀC
IACGA	CCCGRCCRCCR		-		•
5.1	11	21	31	41	- ~
1001 GCCCG	ACGGA TCGCTGA	GATATTCTC	CCATAATGGCA	AAAAAATAGGC TTTTTTTTTCCG	AG TC
CGGGC'	rgccragcgact				
5'	61	71	81	91	
TTTGA	rgtga cctgttt	AGT GTGGCT	CTCCTCTTTTG	AGCATGTGTTA CTCGTACACAAT	3C
1051 AAACT	ACACTGGACAAA	TCA CACCGA	GAGGAGAAAAC	TCGTACACAAT	JG
5 '	11	21	31	41	
<u>۷ ششرن</u> _	י איייייי איי איייייי איי	ATCCAGTGA	ACTCTGCTCTI	CCAAGTGTGTT	CA
1101 TAAAA	TAAA ATATGAG	TAGGTCACT	rgagacgagaa	GGTTCACACAA	ЗT
			81	91	
5'	61	71 TTD GCD CDG(		CTGCACAACGC	СТ
1151 TGTATO	TACCA TCTATATA	AAT CGTGTC	GACGGAAGAC	GACGTGTTGCG	ΞĀ
71021211					
5 '	11	21	31.	41	n
1201 TAGAGA		AATGAGCTTA	AGCT TGTGCTC	TGTTTCTGCTCT ACAAAGACGAGA	7.0
ATCTCI	GGGC CGGAAAG.	IACICGAAT	DADJAJANOJ	110111 HIGHCONGE	
5 '	61	71	81	91	
TTAGGT	CTAAACTATGGT	GT CAGTTTI	AATAGAACAA	AAGTATGCATCT	T
<b>ペーツー カカサククカ</b>	CD $T$ T $T$ C $T$ T $T$ T $T$ C $T$ T $T$ T $T$ C $T$ T $T$	ACAGTCAAA	TTATCTTGTT	TTCATACGTAGA	W

5'	11 TGGCTTGAGCC ACCGAACTCGG	21	. 31	41	
1301 GCCT	TGGCTTGAGCC'	TTTTCGTTTT	CAATG CTGAC	PTCTC CCCT3T(	כיזיכי
CGGA	ACCGAA CTCGG/	AAAAG CAAAA	GTTACGACTG	AAGAGGGGA/A	GAG
5 '	61	71	81	91	
1351 CCTG	TGCTCACCTTA( ACGAGTGGAAT(	CCTTT CCAGA	GTGTA AGGGA	CAACTTTTAAGO	אמני
GGAC	<b>ACGAGTGGAAT(</b>	GAAAGGTCT(	CACATTCCCTC	TTGAAAATTCC	יחירי
	•				
5'	· 11	21	31	41	
1401 CGTG	CCCTGGTAGGG AGGGACCATCCC	GCAT CCCTGT	TCACCAGGTG	CCTGTCATCAC	יכככ
GCAC	AGGGAC CATCCC	CGTAGGGACA	AGTGGTCCAC	GGACAGTAGTC	GGG
		•			
5'	· 61	<b>71</b> ·	81	91	
1451 ACTTO	SACTGA CATCTA CTGACTGTAGAT	CCCTGGTGAC	TATGGGTTCC	TCTTGTTTGTA	CCC
TGAAC	TGACTGTAGAT	GGGA CCACTO	ATACCCAAGG	AGAA CAAACAT	'CCC
	•				-
5 '	11	21	31	41	
AACGG	TGGCT CCAGGT ACCGAGGTCCA	GGAGGCATCA	ATCTGTTGGG	TTCTGGTTCCC	ממר
TTGCC	ACCGAGGTCCA	CCTCCGTAGT	TAGA CAACCC	AAGA CCAAGGG	CCG
	•				
5'	61	71	81	91	
1551 TGCCT	61 TTGGTTTTGAA AACCAAAACTT	AGTCTCTTCT	CTGTATATTC	CTACCCTGCAT	ጥጥር
TEST ACGGA	AACCAAAACTT:	rcagagaaga	GACATATAAG	JATG GGACGTA	AAC
5'	11	21	31	41	•
CTTTG	IGTGG TGCTGAT ACACCACGACTA	rgctgtgcgc:	AGTAGGATTC	TGGATGACTC	דככ
1601 GAAAC	ACACCACGACTA	ACGA CACGCG	CATCCTAAG	ACCTACTGAG	200
		•			
5' ·	61	71	81	91	
1651 ATCAGT	rcacagaetece Agtgtetetgagge	CCTGTTGCA	AGTGTCAGGC	TGACTCGACAC	3TC
TAGTCA	AGTGT CTGAGGG	GGA CAACGTI	TCACAGTCCG	ACTGAGCTGTC	AG
5 <b>'</b>	11	21	31	41	
1701 ACCGTA	AAAT CTGAGTC	AGTCACACAC	AGGCTGTCAG	CCACGGCTTCC	'AC
TGGCAT	AAAT CTGAGTC TTTA GACTCAG	TCAGTGTGTG	TCCGACAGTC	GGTGCCGAAGG	TG
	•				
5'	61	. <b>71</b> .	81	91	
1751 TTGCAT	GGCTATTCTAT CCGATAAGATA	TTTCACACGT	GAGTTTCTGT	TGCTGGCTGGC	TG
AACGTA	CCGA TAAGATAI	<b>AAAGTGTGCA</b>	CTCAAAGACA	ACGACCGACCG	AC
5'	11	21	31	41	
1801 ACTGGC	atta tetatge:	'AAGTTGAAA	TCAGGAGTGC	CCAGCAGAGCC	CA
TGACCG	TAATAGATACGA	ATT CAACTTT	AGT CCTCACG(	GT CGT CT CGG	GT
•		•			
5'	61	71	81	91	
1851 TCATTCT	CAC TGTCTTTC	SAA ACAAAGC	TTTDDDATD	SATCGATGAAC	ЗT
AGTAAGA	rcac tgtctttg Agtgacagaaac	TTTGTTTCGA	CATGCCAAAC	TAGCTACTTG	ZĀ.
5'	11	21	31	41	
1901 ATTTAAA	GCATTTCATGC	AATGACAAAG	TGCTCAGTAG	TGGAAGGCAGC	3C
שירויים על עלים ביי	יכים את את מיים איני	നന മ വനവനന്നാ	IN CICIN CICIN THE	12 0 000000	

`5 ¹		71	81	91	
1951 TG	TGACCAGT CTGCCT( ACTGGTCA GACGGA(	SCTCCTTAC CGAGGAATG	PATAATTGTGA( ATATTAACACT(	GGATTTGTTACTG CCTAAACAATGAC	G
5'.	11	21	31	41.	
2001 AA	CAGTACATGGAGGCC GTCATGTACCTCCGC	CTGA CCTTG: SACTGGAAC	PGGGGGCACAG( ACCCCCGTGTC	GTGGAACCTTAG CCACCTTGGAATC	C G
. 5 '	61	71	81	91	
2051 AC	AATATAGTGTGTGTC CTATATCA CACACAC	AGTTCTCC	TCAGTCAGGGTAC	SATCGAGTCACGA	G
5'	11		31		_
2101 AA	CTCCAGG TACTATA AGAGGTCCATGATAT			•	A I
5'	61	71	81	91	
2151 TAT	ATCCCCAAACACTT TTAGGGGTTTGTGAA	GTTTATCGT CAAATAGCA	GTAG CGTACCT CATCGCATGGA	'AAAAGACTATTC' LTTTTCTGATAAG	Г А
5 '	11	21	31	41	
2201 ATT	TATGGGTGTCCCCAC ATACCCACAGGGGTG	TTT CTTGGT AAA GAACCA	TTGG TCACCCC AACCAGTGGGG	GAT CCCCCGGTC: CTAGGGGGCCAG	r A
5 '	61	71	81	91	
2251 TCT AGA	GCTGTATCTAGAAC CGACATAGATCTTG	<b>Α</b> ΥΤΟΑΡΤΆΛ	AAATGATGTAT	GGGAATAGTGTT	r
5'	11	21	31	41	
2301 CCA	TATGATCTGTTGTC ATACTAGACAACAG				L A
5 '	61	71	81	91	
2351 AAA	ACCCAGTGCAGCTGA TGGGTCACGTCGACT	ATGATGCAA TACTACGTT	AGCA GTCTCTC TCGT CAGAGAG	TCTGTGTACAGT( AGACACATGTCA(	3 7
5 '	11	21	31	41	
2401 CCC GGG	CACCTATTTAAAAA GTGGATAAATTTTT	CA CGTACA GTGCATGT	ASCC CAGAACA ISGG GTCTTGT	CTGTGAAACACTT GACACTTTGTGAA	7
5 '	61		. 81	91	
2451 AAC	ATAAGAA CAAACGCA PATTCTTGTTTGCGT	GCGTCTGG CGCAGACC	ATTCTTTCCAA TTGGAAAGGTT	GGAGAGCAGCTTT CCTCTCGTCGAAA	7
5 '	11	21	31	41	
2501 CTCC	CACAGGAACACAGTA STGTCCTTGTGTCAT	ACAAAAGA( TGTTTTCT(	GTC CGCCGCC CAG GCGGCGG	ATCCACACCCAGC PAGGTGTGGGTCG	•
5' ·	61	71	81	91	
2551 CAAC	SACACCT CAGAGGCC	ATAGGGACA	ACCTCCTTGCT	CCGCTTGTGGA	ļ

5 '	11	. 21	31	41	
2601 CTG	GAGCAGGGGCAC	AGGTC CCAGC	AACTGATCCTC	AGTG GATGGGT	C
ZOOT GAC	GAGCAGGGGCAC CTCGTCCCCGTG	TCCAGGGTCG	TTGACTAGGAG	TCACCTACCCAC	3G
		•			
5'	61	71	81	91	
2651 CAG	TCAAAGCCTTAA: AGTTTCGGAATT!	TGGGCTCTCT	TTTGAAGGGGA	AAGAAAGAATTT	.'C.
5 '	TTATGATATCCAA AATACTATAGGTT	21	31	. 41	
AGC	TATGATATCCA	CATTATTAT	AGTTGATGAGT	፲፰፫. ፲፮ <b>፫</b> ፻፮፮፮፻፹፫ሮአ	. 7. 2
2701 TCG	AATACTA TAGGTI	GTAA TAATA'	ICAACTACTCA	ATCATTTAAGGT	مكر
5'	61	, <b>71</b>	81.	91:	
2751 AAAA	AAAGATGATTTT TTTCTACTAAAA	'ATATGTATG	CATAAAAAAA	ATCTTTGTAAAG	TC
- · TTTT	TTTCTACTAAAA	TATACATACT	GTATTTTTTT	TAGAAACATTTC	AC
5 '	7 7	21	Эл	49 .	
	አርጥርርስ ልጥስ አጥጥ	ሬ ደ ጥክ ክ አ ረ አ ረ ረ ጥረ	anany an camanaca T	41	
2801 GCGT	AGTGCAATAATT TCACGTTATTAA	ATTTCTCCAC	LIAICIIIGCA LAATAGAAACGT	ים א אייייייתיית א מ' ג' ג' ב' LT T ATAAATTA	TA
			millomuco.	www.rut.t.WWI	H.I
5'	61	71	81	91	
2051 AATA	TTGTACATGTGT( AACATGTACACA(	GTAA TTTTTC	ATGTATTCATT	TGCAGTCTTTG:	ГА
ZOSI TTAT.	AACATG TACACA(	CATTAAAAAG	TACA TAAGTAA	ACGTCAGAAAC	AΤ
5 '	11	21	31	41	
2901 אמת ב	AAAAAA CTTTACT FTTTTT GAAATGA	CAAMACAAA	GTATAATAGAA Camammamam	CATTAATCATT	ľA
Ann.	IIIIII GAAAIGA	CHAIACHAA	CAIAIIAICII	GIAATTAGTAAA	7.T.
5 1	61	71	81	91	
OCT TTATA	ACTCA GACAAGG	TGTAAATAA	ATTCATAATTC	AAA CAGCCAGTZ	ıт
AATAT	ACTCA GACAAGG TTGAGT CTGTTCC	ACATTTATT	TAAGTATTAAG'	TTTGTCGGTCAT	À.
	•				
5'	11	21	31	41	
3001 ATATO	CATATATGGGTG GTATATACCCAC	TTA CATTGC	AAAAATCTCTA:	CTTTGTTCTAT	T.
		AAI GIAACGI	TTTTAGAGATA	AGAAACAAGATA	A
5'	61	71	81	91	
CACAT	GCTTAAAGAAGT	AAGAAATCTT	TTGTGGATATG	ጉር ምልጋልምምልምል ልሞ\$	יעי
3051 GTGTA	GCTTAAAGAAGT CGAATTTCTTCA	TTCTTTAGAA	AACACCTATAC	ATTAATATGTA	T
					-
5 '	11	21	31	41	
3101 TAAAG	TATATATATATG	TATGATACAT	GAAATATATTI	'AGAAATGTTCA	T
ATTTC	ATATATATATACA	ATACTATGTA	CTTTATATAAA	TCTTTACAAGT.	A
5 '	61 ·	<b>71</b> .	81	91	
_					_
3151 TTAAA	TAATGGATATTCT ATTACCTATAAGA	AACCACACT	ጉዶጥ፣ ፲ይደረ፤ H የልጥጥልልርጥጥልጥ	CTTCTD N N N N N N N N N N N N N N N N N N N	: <u>\</u> ኮ
				+ OIRRAMI	-
5'	11	21	31	41	
3201 AATGAA	<mark>፞</mark> ፟፟፟፟፟፟፟፟፟፟፟፟፟፟፟፟፟፟፟፟፟፟፟፟፟፟፟፟፟፟፟፟፟፟፟፟	<b>XAAAAAAAA</b>	AAAAAAA		
mmn x mmm	ארוארדו נדארדוניו ולוינדייו או ולוינדארדוא		WINTER CONTROL		



## Figure 12

5 '	. 1	1	21	31	41
	1 Aagtgtaaata	AAATAAACA	TCTAATAAA	<b>LAAAATTACAT</b>	ACCATAGAGG
•	<sup>L</sup> TTCACATTTAT	TTTATTTGT	<b>AGATTATTT</b>	TTTTAATGT?	TGGTATCTCC
5 '	6	1	71	81	91 .
	AACAAGATAAT TTGTTCTATTA	TTCTGCCCA	ACTTCATACO	CTCCAGCGTA	TAGTGTTGAG
5:	TTGTTCTATTA	AAGACGGGT	TGAAGTATGG	GAGGTCGCAI	ATCACAACTC
		•		•	
5.1	1	1	21	31	41
<b>J</b> .	ርጥጥጥር/ርጥርጥር ጥ	ͳϲϹͲϲͲϲͲΔ	ጥጥርጥል ልጥርነጥ	ATGTTAAATI	CTCTACCTGA
103	CAAACCAGACA	ACGACACAT	דמטמידימטממ	TACAATTTAA	GAGATGGACT
	CAMACCAOACI				
5 '	6.	L	71 <sup>.</sup>	81	91
5		ב. האמא אמשמא	᠈᠇ᠳᢕᡎ᠘᠕ᡎ᠘ᡆ	ייתייאייאמאמייי	T ጋንሆን ሞተርሞር Δ
153	AGGTCTAGGCC'		TA A CA CTA CA		AACAACACGT
	TCCAGATCCGG	AIGIICACI	IMAGAGIACA	MAINICI CE	22104101002
	4 *		<b>0.1</b>	21	<b>4</b> 1
5 '	AACCTTGTTCC	L .	7 T		7.7 7 7 7 7 7 7 7 7 7 7 7 7 7 7 7 7 7 7
201	AACCTTGTTCC	L'IAA'I"I'AA	AACTATGGTT		MCHAHAHCI GG
201	TTGGAACAAGG	ATTAAATT'	ITGATACCAA	A.L.L.I.I.G.T.T.	IGITITGACC
				0.4	91
5 '	61		71	81	
251	CTACAGCCAATA	ACTGAAGG	<b>EGGTTACCTT</b>	GTTGAAGGGG	TGGAAAAGAG
251	GATGTCGGTTAT	TGACTTCC	CCCAATGGAA	CAACTTCCCC	ACCTTTTCTC
	•				
5 '	11	. 2		31	41
207	AGAGGAGGAAGA TCTCCTCCTTCT	AGGGAGTT	CAAGAGAAGG	AGAAGAACAA	GAGGAGAGGA
301	TCTCCTCCTTCI	TCCCTCAAC	TTCTCTTCC	TCTTCTTGTT	CTCCTCTCCT
5'	61	. 7	71	81	91
	GGAAGCTGCCAC	GAGGGGAG	TGGGCCATG	AGAACTTGGC	CAGGAGAAAT
351	GGAAGCTGCCAC CCTTCGACGGTG	CTCCCCTCI	ACCCGGTAC	TCTTGAACCG	GTCCTCTTTA
	•	•			•
5 '	11	` 2	1		41
	AGCCAGTATETG	GAGTACACC	ACTGAGGAG	GTAGCCAGGC	TAGCAGTTAG
401	TCGGTCATAGAC	CTCATGTGG	TGACTCCTC	CATCGGTCCG.	ATCGTCAATC
				•	
5 1	61	7	1	81	91 ·
_	AAGAGTAGATTA TTCTCATCTAAT	GGGGTTATT	TTTCCCCCA	CTCCACATAG	TTATCAAAGC
451	TTCTCATCTA AT	CCCCAATAA	AAAGGGGGT	GAGGTGTATC.	AATAGTTTCG
51	11	2	1	31	41
J	 \_ \_ \_ \_ \_ \_ \_ \_ \_ \_ \_ \_ \_ \	 ייים אים העדירייי	_ GAGTCTCAT(	CTATTTGTAA	GCTAGTTGGG
501	CAAATAAAATAA GTTTATTTTATT	CCATATCAGA	CTCAGAGTA	GATAAACATT	CGATCAACCC
	GILIMITIMIA	301711 01011		<del></del>	
5 '	61	7	1 4	81	91
<b>5</b> .	 המע ענהמוא ארא ענה אות ע	י מהלים לי ליום לי			
551	TATAAGATTAAT ATATTCTAATTA	T T GGCTGTW	CIACAGIII	TCTAAAGATT	TATCCTTGA
•	WIWIICIWWI IW	MCCOMCH I	GALGI CAAA.	CIMMICILLY	
<b>~</b> .	11	2:	<b>.</b> .	31	41
5'	II ATCAAAAACTTGO				
601	ATCAAAAACTTG(		344641 <b>3</b> 616	ንጭርማጥ አጥ ተ ተ ተ . የ ለ ለ ለጥ አጥ ተ ተ ተ .	A A TO A THURTH

5 '	61	71	8 <b>i</b>	91	
651 ATT	TATATTG TTTGCA ATATAAC AAACGT	CTTTCTAAA	GTTTCTTCTAA	ATGTTCCATG	GTC
AAT	ATATAACAAACGT	GAAAGATTT	CAAAGAAGA'I"I	TACAAGGTAC	CAG
5'	11	21	31	41	
ZOI AAT	TAAAAAA TATACA ATTTTTTATATGT	TATTGGCTA'	TTAAA TTCGTC	TAAG TGGGGC	TGG
TTA	ATTTTTTATATGT.	ATAA CCGATI	AATTTAAGCAG	ATTCACCCCG.	ACCI
5'	61	71	81	91	
GAG	ATAGCTCAGAGGT	TAAGAGCAC	rgactgctctt	CCAGAGGTCC	TGAG
751 CTC	ATAGCTCAGAGGT PATCGAGTCTCCA	ATTCTCGTG	ACTGA CGAGAA	GTCTCCAGG	ACTO
		2.5			
5 ' mmax	11	21	31	41 האמות מות בייבים	2 m 2 C
801 1102	ATTCCCAGCGAC TAAGGGTCGCTG	LACA IGGIGG	GAGTGTCGCCA.	CIGIAATAGA AGACATTATAGA	ያችኒዎር ሊኪዊብ
71102	1111000100				****
5 '			81		
851 GATO	TGACGC CCTCTTC ACTGCG GGAGAA	TGGAGTGTC	TGAAGACAGCT	'ACAATGTAC'	CAT
CTAG	ACTGCG GGAGAAG	ACCTCACAG	ACTICICICE	TGT TACATG	4GTA
	11	21	31	41	
OOT ATAT	AATAAAT AAATTA' TTATTTATTAAT	TAT TAGAAA	ATTCTTCTAAG	TGTATCATT	(ATA
JOT TATA	TAATTTATTAAT.	'ATAATCTTT	TAAGAAGATTC	'ACA TAGTAAI	TAT
5'	61	71	81	91	
GAAT	ATTTAA TATATAA	AGTAAATGC	CTCAGGAAATA	TAAACTTGG	ATT
951 CTTA	ATTTAA TATATAA TAAATTATATATT	TCATTTACG	GAGTCCTTTAT	ATTTGAACCT	TAA
- 1	11	0.7		47	
5 ' 777 T	ርአአአርአአርጥጥርአጥ LL		. 31 3CCD CDDDDDD	ቸች ተርፈርሲያ ርርያር	ccc
1001 TTTA	CAAAGA ACTTCAT GTTTCT TGAAGTA	CTCATCACC	CGGTGTTTTT	ACACATGGTO	CCC:
5'		71	81	91	:
1051 AAGA	CCGGAG GGAGGGG GCCTC CCTCCCC	AGAAGGAAG(	GATGGAGATA	GAATTTTGCC	TCT
			CIACCICIAI	CIIAAAACGG	MOM
5 <b>'</b>	11	21	31	41	
1101 GCAT	rccttgggctggcz	ACAGGTATA	TGCTGTGGGA	ATTGGGAAAC	TAC
CGTAI	AGGAAC CCGACCGT	"GT CCATAT"	'ACGACACCCI'	PAA CCCTTTG	ATG
5 '	61	71	81	91	
	AGCTG CAAAGCTG TCGACGTTTCGAC	GG CGGAACI	CGTTTCCGCA	AGCTGGGCTC	ATC
TIST TTCCT	TCGACGTTTCGAC	CCGCCTTGA	GCAAAGGCGT	rcgacccgag	TAG
5 '	11	21 ·	31	41	
					TGT
1201 ATTCA	GTCCATGCATGGC CAGGTACGTACCG	ACGGTGTGA	CGTCACTTGAZ	ATTTTGTAA	ACA
	<i>C</i> 7	<b>.</b> .	.0.4	0.1	
5'	61	71 ଫଟ ଫଟର ପର ର	<sup>.</sup> 81 ຫາວເຫລເເລລລອເເ	91	aam
1251 GTTCC	AGAGA TGTAGAGA TCTCTACATCTCT	ACCACAA ACCAGTGTT	AUC AUCHAAGC	;GCCCTCCCT	GCD.

5'	11	21	31	41	
1301 ATT	rccagactaaga Aggtctgattct	ggaag aaaa <i>a</i>	CCATTGCTGAT	TAAA CATCTGC	λT.
TAA	AGGTCTGATTCT	CCTTC TTTT1	GGTAA CGACTA	ATTTGTAGACG1	'A'
E 1	61	71	0.7		
יות אל	61 SCGCCCCACCTC CGCGGGGGTGGAC	/ L ጉርካጥአ ርካርካር	יז כיז כיז כיז כיז ביז כיז כיז כיז כיז כיז כיז	91	~-
1351 ACT	CGCGGGGGTGGAG	GTATGTGTG	TGTGTGTGTGTGT	CACACACACACACACA CTCTCTCTCTCTCT	C
			10101010101	010101010101	G.
5 '	11	21	<b>31</b> ,	41	
1401 CAAC	CAAACAGAACAA GTTTGTCTTGTT	ATACACATG	CATGT CTACAG	CCTG CAGGAACA	AZ
TTO GTTO	GTTTGT CTTGTT	TATGTGTAC	GTACAGATGTC	GGACGTCCTTGT	ΤŢ
<b>E</b> I	61	71	0.7	0.7	
אייטילי אייטילי	™₮₼₼₼₼ ₽₩	/ T .	81	91	
1451 TACC	TATGTCTGTGAG ATACAGACACTC	CHAC CAGGA CTTCCTCCT	CATGUACAGGT CTACCTCTCCA	CCTAACCTCTGT	CI
	•			•	
5'	11 CAAGCC CTGAAG GTTCGG GACTTC	21	31	41	
1501 CCTA	CAAGCC CTGAAG	TCTGGTCAG	GTCAAATGTA	CAAAAGCAGGCT	AA
ISUI GGAT	GTTCGGGACTTC:	AGAC CAGTC	CCAGTTTACATO	STTTTCGTCCGA'	ГT
	<i>C</i> 1				
5'	ianammyanay ot	/ L .	81	91	
1551 COTT	GCTGTTTAGTGA CGACAAATCACT	HAGAIIIII CTCT DD X X X X X	AGN A GTTGAGT	CAGGAACAACCTI VTCCTTCTTCCT	ων F.T.
CCII	concumia chei.	LICIMMM	MONNGIIGAGA	icciigiigga.	Ι₩
5' ·	11	21	31	41	
1601 TTCC	TAGGAT TTGGAGA ATCCTAAACCTCT	GTG CTCAGG	AGGAAACATTC	AGA CAACTGATO	ЗC
AAGG/	ATCCTAAACCTCT	CACGAGTCC	TCCTTTGTAAG	TCTGTTGACTA	]G
5 '	61	<b>77 7</b>	81		
	יכיזיכיזיא פיבי יכיזיכיזיא פיביריא פא	. ጥጥር እርደጥእጥ . ሊጥር እርደጥእጥ	™CCCC™™C™™™ O T	CALCACA AC	чт
1651 AGAGA	GTGTA CCCCAGA CACAT GGGGTCT	'AAGTCCATA	ACCCCATCAAT	CAA CACGAGTAC	Δ, 1.
					~ .
5 '	11	21	31	41	
1701 ATGTG	CTAGA TATATTA GATCTATATAAT	GCA CAGCCT	GCCTTCTGCTG	CACAACGCCTTA	'G
TACAC	GATCTATATAAT	CGTGTCGGA	CGGAAGACGAC	GTGTTGCGGAAT	'C
5 '	61	71	81	01	
AGACC	CGGCCTTTTCAAT	GAGCTTAGC	₽₽₽₽₽₽₽₽₽₽₽₽₽₽₽₽₽₽₽₽₽₽₽₽₽₽₽₽₽₽₽₽₽₽₽₽₽₽₽	ᠠᠬᡳᡙᢗᡊᡙᡳᡎᡳ ᡓᢧ	Δ.
1751 TCTGG	CGGCCTTTCAAT( GCCGGAAAGTTA(	CTCGAATCG	AACACGAGACA	AAGACGAGAGAA	T
5'	11 AAACTATGGTGTC ITTGATACCACAC	21	31	41	
1801 GGTCT	AAACTATGGTGT	CAGTTTTAAI	'AGAACAAAAG'	<b>FATGCATCTTGC</b>	C
CCAGA	I'I'IGA TACCACAC	FTCAAAATTA	TCTTGTTTTC	ATA CGTAGAACG	G
5 '	61	71	81	91	
	TGAGCCTTTTCG	TTTTCAATG	ያርጥር <u>ልር</u> ጥጥርጥር	, הכה שינהנה התהרבה הבה	т
1851 AACCGA	TTGAG CCTTTTCG AACTCGGAAAAGC	AAAAGTTAC	GACTGAAGAG	GGAAAGAGAGG	Ā
					_
5'	11	21	31	41	
1901 GTGCTC	'ACCTTACCTTTC 'TGGAATGGAAAG	CAGAGTGTA GTCTCACAT	AGGGACAACTI TCCCTCTTGD D	TTAAGGAGGCGT	Ċ

5 '	61	<b>71</b> .	81	91	
1951 GTCC	CTGGTA GGGGCA GACCAT CCCCGT	TCCCTGTTC	ACCAGGTGCCT	GTCATCACCCC	\CI
CAGG	GACCATCCCCGT	'AGGGACAAG'	rggtccacgga	CAGTAGTGGGG.	ľGA
5 '	11	21	31	41	
ጥርልሮ	TGACAT CTACCC	.ፚፗ ጥርርጥርክርጥስባ	₽ĊĊĊ₩₩ĊĊ₩Ċ₩ ₽¥	ჀႺჀჀჀႺჀਸ਼ <i>ੑ</i> ੑੑੑੑੑੑੑਜ਼ਸ਼ੑੑੑਜ਼ੑੑੑੑਲ਼ੑੑਜ਼ਸ਼ੑੑੑੑੑਜ਼ਸ਼ੑੑੑਜ਼ਸ਼ੑੑਜ਼ਸ਼ੑੑ	
2001 ACTG	ACTGTA GATGGG	ACCA CTGATA	CCCD AGGAGA	1G111G1AGGGA	
11010.				"Canification of the	. 10
5'	61	71	81	91	
GGTG	GCTCCAGGTGGA CGAGGTCCACCT	GGCATCAATC	TGTTGGGTTC	TGGTTCCCGGCT	GC
ZUSI CCAC	CGAGGT CCACCT	CCGTAGTTAG	ACAA CCCAAG	ACCA AGGGCCGA	<b>LCG</b>
5'	11	21	31	41	
2101 (7777)	GTTTTGAAAGT CCAAAA CTTTCA	CTCTTCTCTG	TATATTCCTA	CCCTGCATTTGC	TT.
GAAAG	CAAAACIIICA	SAGAAGAGAC	ATATAAGGAT	GGGA CGTAAACG	ΑA
5'	61 CGTGCTGATGC CCACGACTACGA	71	81	91	
ጥርጥርባ	rGGTGCTGATGC	רכיזכ רכיר <i>א</i> כר	ልርር <b>ል ጥ</b> ጥርጥጥር	ユカጥር <b>カ</b> ርጥርጥርር አ	TC
2151 ACACA	CCACGACTACGA	ACACGCGTCG	TCCTAAGAAC	CTACTGAGAGGT	'AG
<b>5¹</b> .		21	31	41	
AGTCA	CAGACTCCCCCT GTCTGAGGGGGA	GTTGCAAAG	TGTCAGGCTG	ACTCGACAGTCA	CC
TCAGI	GTCTGAGGGGGI	CAA CGTTTC	ACAGTCCGAC.	rgag ctgtcagt	GG
	·~~		•		
5'	61		81	91	
2251 GTAAA	ATCTGAGTCAGT TAGACTCAGTCA	CACACACAG	GCTGTCAGCCA	ACGG CTTCCACT	TG
			•		
5 (	11 CTATTCTATTT GATAAGATAAAA	21	31	41	
CATGG	CTATTCTATTTT	CACACGTGA	GTTT CTGTTGC	TGG CTGGCTGA	CT
2301 GTACC	GATAAGATAAAA	GTGTGCACT	CAAAGACAACG	ACCGACCGACT	GA
5 ¹ . ·		71	81 :	91	
2351 GGCAT	TATCTATGCTAA ATAGATACGATT	GTTGAAATC	AGGGGTGCCCA	GCAGAGCCCAT	CA
CCGTA	ATAGATACGATT	Caaci'i'iagi	.CCCCACGGG1	CGTCTCGGGTA	₹.T.
·5 t	11 '	21	31	41	
. מאניהוניה.	ልርጥርጥርጥጥፕሮልል ነ	ልሮል ልልፎርጥረባ	<sup>∙</sup> ∆℃₢₵₸₸₸₢₯₸	ም ነው። የርያ ጥርያ አርርጥል፣	ילינו
2401 AAGAG	ACTGTCTTTGAA IGACAGAAACTT	TGTTTCGACA	TGCCAAACTA	GCTACTTGCATA	ĀĀ
5 '	61	71	81	91	
2451 TAAAGO	CATTT CATGCAAT STAAAGTACGTT <i>I</i>	rga caaagtg	CTCAGTAGTG	GAAGGCAGGCT	ΞT
ATTTC	TAAAGTACGTTI	ACTGTTTCAC	GAG TCATCAC	CTTCCGTCCGA	<u> </u>
<b>-</b> •	9.9	0.1		. 49	
51 .	11	21	31 <sub>.</sub>	41 mm/mm//mm//	. ~
2501 GACCAG	TCTGCCTGCTCC AGACGGACGAGG	TTACTATAA	T T.G. T.G.A.G.GAT.	T T G T.T.W.C.T.G.C.W.	IC.
C16610	OOMOOMOO	mm IGMIMI I.	WWCWCICCIW	AACAAIGACCII	ی.
5'	61	71	81	91	
	TGGAGGCCTGAC				<b>Δ</b> !
2551 man mam	A COTT CCCC A CITIC	Ch handaaa	COT COT COT CO	TOTAL STREET	•

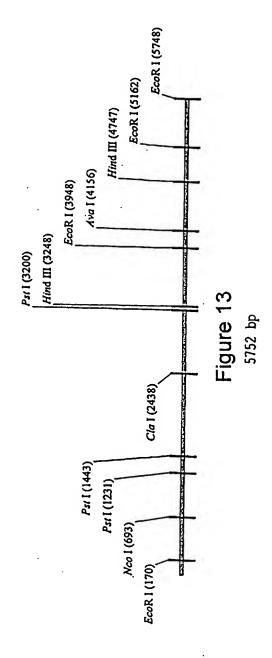
5 '	11	21	31	41	
2601 ATA	TAGTGTGTGTCT( ATCACACACAGA(	CAAGAGGAAG'	<b>CAGGGTACTA</b>	GCTCAGTGCTCA	A.
TAT	ATCACACACAGA(	STTCTCCTTC	agtcc catgai	CGAG TCACGAGT	T
5 '					
		71	81	91	
2651	CAGGTAC TATATA GTCCATG ATATAT	ATACATTTGC	CCGTTTTATCT	'CTAA TGTGAAA'I	ΑZ
GAG	GICCATGATATAT	ATGTAAACG	GCAA AATAGA	GATTACACTTTA	T
<b>5</b> • .	77	21	2 4	4 7	
ATTC	11 CCCAAACACTTGT GGGTTTGTGAACA	עייייט יוי כעימעיאמא ד ס	. C.C.C.C.D.D.D.D.	41	
2701 TAG	CCCAAACACIIGI CCCTTTCTCIII	TIMI CGIGIA	CCCD TACCTAA	AAGACTATTCTA	TT
		MINGCACAL	CGCAIGGAII	IICI GAIAAGAT	AA
5 ¹	<b>61</b>	71	81	91	
ATG	GTGTCCCCACTT CCACAGGGGTGAA	TCTTGGTTTG	ייייייייייייייייייייייייייייייייייייי	<b>ふしているしょうしょうしょうしょうしょうしょうしょうしょうしょく しょうしょう ストップ・スティー・フェー・フェー・ファー・フェー・ファー・ファー・ファー・ファー・ファー・ファー・ファー・ファー・ファー・ファ</b>	CIT
2751 TAC	CACAGGGGTGAA	AGAACCAAAC	CAGTGGGGCT	AGGGGGGCCAGAA	GD CI
			•		
5'	11	21	31	41	
GCTG	11 TATCTAGAACAG ATAGATCTTGTC	TGACTATAAA	TGATGTATGG	GAATAGTGTTTC	CA
ZSU1 CGAC	'ATAGAT CTTGTC	ACTGATATTT.	ACTA CATACCO	CTTATCACAAAG	GT
5'	61	71	81	91	
2851 TATG	ATCTGTTGTCTG	GAGTATATGC'	TACA TGTTCA?	CTTA CTGTACAA	AA
ATAC	TAGACAACAGAC	CTCATATACG	ATGTACAAGTA	AAATGACATGTT:	ГT
e i	4 4	· ·	2.5		
5'	TT	21 21	31	41	
2901 ACCC	11 AGTGCA GCTGATO TCACGT CGACTAO	SATG CAAAGCA	AGTOTOTOTOT	GTGTACAGTGC	3C
1.666	ICACGI CGACIAC	TACGITICG	CAGAGAGAGA	CACATGTCACGC	зG
5 '	61	. 71	81	97	
CACC	TATTTA AAAATCA	CGTACTTGCC	TCAGAACACTG	ייעבע אטרטריייט ז מעריייט אטרטרייט זי	\ C
2951 GTGG	61 TATTTAAAAATCA ATAAATTTTTAGT	GCATGAACGO	GTCTTGTGAC	ACTTTGTGAATT	iG.
5,	11	21	31	41	
3001 ATAAC	GAACAA ACGCAGC	GTCTGGATTC	TTTCCAAGGA	GAG CAGCTTTCI	'C
TATT	11 SAACAA ACGCAGC CTTGTTTGCGTCG	CAGACCTAAG	AAAGGTTCCT	CTCGTCGAAAGA	ď
5'	61	71	81	91	
3051 CACAC	61 GGAACA CAGTAAC CTTGTGTCATTG	AAAAGAGGTC	CGCCGCCATC	CACACCCAGCCA	A
GIGIC	CIIGIGICATIG	TTTTCTCCAG	GCGGCGGTAG	GIGIGGGICGGI	T.
5 (	. 11	21	31	41	
GACAC	ירידראה אהכהרראיזי ירידראה אהההרראיזי	AGGGACAACC	₽₽₽₽₽₽₽₽₽₽₽₽₽₽₽₽₽₽₽₽₽₽₽₽₽₽₽₽₽₽₽₽₽₽₽₽₽	ች ፲ ግርኒክ አርኒክ ርርጥርርጥ	C
3101 CTGTG	CTCAGAGGCCATA GAGTCTCCGGTAT	CCCTGTTGG	AGGAACGACC	CAACACCIGCI CATTGTGGACGA	C
					_
5 🗥	61	71	81·	91	
3151 GAGCA	GGGGCACAGGTCC	CAGCAACTG	ATCCTCAGTGG	ATGGGTCTGCA	G
CTCGT	GGGGCACAGGTCC CCCCGTGTCCAGG	GTCGTTGACT	RAGGAGTCACC	CTA CCCAGACGT	C
5'	11	21	31	41	
3201 CCAAA	GCCTTAATGGGCT	CT CTTTTGAZ	AGG GGAAAGAA CCCCCCCCCCCCCCCCCCCCCCCCCCC	AGAATTTCAAG	7 - 7

5 '		61	71	81	91
5	ምምስም <b>ር</b> ልሞልሞር	CAATATTATT	TATAGTTGATO	AGTTAGTAAA	TTCCAAAAAA
3251	TTATGATATC AATACTATAG	GTTATAATA	TATCAACTAC	TCAATCATTI	AAGGTTTTTT
	1111110111111				
5 '		11	<b>21</b>	31	41
	AAAAGATGAT	TTTATATGT	TGACATAAAA	AAAATCTTTG	TAAAGTGCGC
3301	AAAAGATGAT TTTTCTACTA	AAATATACAI	ACTGTATTT	TTTTAGAAAC	ATTTCACGCG
				•	01
5 '		61	71	81	ን <b>፲</b>
2251	AAGTGCAATA	ATTTAAAGAG	GTCTTATCTT	TGCATITATA	ለጥጥ የሚተፈጥር የጥጥ፣
335T	AAGTGCAATA TTCACGTTAT	TAAATTTCTC	CAGAATAGAA	ACGIAMAINI	IIMMINIIIM
			21	31	41
5'	ATTGTACATG	11	ייים עניייטים ערייים	ኯጋልግጋጕጕጥለላጉ፣	יבידים בידים
3401	ATTGTACATG TAACATGTAC	TGTGTAATTI	TTCAIGIALI	GTAAACGTCA	GAAACATAAA
3401	TAACATGTAC	ACACATTAAA	MAGIACAIAA	01/11/100	
		61	71	81	91
5 '		61 macricarnamo	ኒጥጥርጥልጥል የልልጥልጥል	AGAACATTAA	TCATTTATTA
3451	AAAAAAACTT TTTTTTGAA	ATCACAATAC	AAACATATTA	TCTTGTAATI	'AGTAAATAAT
	TTTTTTGAA	AIGACAAIAC			
- 1		11	21	31	41
<b>5</b> .	TAACTCAGAC ATTGAGTCTG	AAGGTGTAAA	TAAATTCATA	LATTCAAACAG	CCAGTATATA
3501	ATTCACTCTG	TTCCACATTT	ATTTAAGTAT	TAAGTTTGTC	GGTCATATAT
	MITOMOTOTO				
51		61	71 .	81	91
	TGCATATATG ACGTATATAC	GGTGTTACAT	TGCAAAAATC	TCTATCTTTG	TTCTATTCAC
3551	ACGTATATAC	CCACAATGTA	ACGTTTTTAG	AGATAGAAAC	AAGATAAGTG
				31	41
5'		11	21	ጋፗ እመኔሙረመኔንምጥ	
2601	ATGCTTAAAG.	AAGTAAGAAA	TCTTTTGTGG	MIAIGIAAI I MUTALGIAAI I	TATGTATATT
3601	TACGAATTTC'				
		· ·	71	81	91
5'		▄ ▄ ▄	$\lambda \subset \lambda \cap C \cap D \cap D \cap D$	TATTTAGAAA	TGTTCATAAT
3651	AGTATATATA TCATATATATATATATATATATATATATATA	LWIGIWIGWI VANCALWIGWI	TCTACTTTAT	ATAAATCTTT	ACAAGTATTA
51	:	1.1	21	31	41
5					TTTTTAAAAT
3701	TTTAATGGAT! AAATTACCTA!	TAAGAAACCA	CACTTATTAA	CTTATGTTGT	AAAAATTTTA
	Mult 11100111				
5'	6	51	71	81	91
	) KAAAAAAAAA	AAAAAAAAA.	AAAAAAAAAA	YAYAT.I.I.I.I.I.I.	
3751	AAAAAAAAAA TTTTTTTTTT	TTTTTTTTTT	TTTTTTTTTTT	.I.I.I.I.WAAAAA	Winner
	_		<b>0.</b> 7	31	41
5 '	t TTATTCCAGA	11	21 ************************************	ታቷ ጥልልሮሮጥፕሮልል	
rose	TTATTCCAGA ( AATAAGGTCT (	JATTAAAGAC.	MCTAGATCTT	ATTGGAACTT	CCCGTCCGTT
	AATAAGGTCTC	TAMITICIG	IGMICINGAM		
е.	4	i1 '	71	81	91
ອ '	GAGGTCGGCAA	/ ፲ ለጥርረርጥርጥሮ ፮ ፮ (	CATAGAAGTC	AGGGACCATT	TTCTTCTTGA
รครา '	GHOG I COGCHY		CHAMCHTCAC	ѵҁҁѽӆѽҾѦѦ	AAGAAGAACT

5 '		11	21	31	41
3901	ACATGCAGT	CACTTTCCTG	ATTGCTCTTC	LA CATCCTCAA:	GGCTCCGGAAT CCGAGGCCTTA
J J U I	TGTACGTCAC	TGAAAGGAC	TAACGAGAAG	I GIAGGAGI I	CCGAGGCCIIA
. <b>5</b> I		<b>61</b> .	71	81	91
2051	TCCGGGGGT	TGGTGGGCT	TTGATCTCAG	GACTCTGGAG	G CAGAAGCAGG
3951	AGGCCCCAC	CACCACCCGA	AACTAGAGTC	CTGAGACCTC	CGTCTTCGTCC
٠.		11	21	31	41
5'	CACATCTCTC				AGCTCCAGACC
4001	GTCTAGAGAC	ACTTATACT	CCGGTCGGAC	GTGATGTGTC'	TCGAGGTCTGG
•					<u> </u>
5 '		61	71	81	91 • • • • • • • • • • • • • • • • • • •
4051	AGTCATGGCT	'ACATCATGA!	AACCCTGTCT rrcccacacaca	С <del>ДААААСААА</del> СФФФФФФ	ATAAAAACTGT FATTTTTGACA
	TCAGTACCGA	ITGIAGIACI	LIGGGACAGA		i i i i i i i i i i i i i i i i i i i
51		11	21	31	41
	TGTGTTTCTA	CCATAGTGTT	TAAACTCAGA	GTCTGAGTAA'	rgtcgggctga
4101	ACACAAAGAT	'GGTATCACA!	ATTTGAGTCT	ÇAGACTCATT	ACAGCCCGACT
		61	71	81	91
5'	CA TO COTO COCO	תכיויתים א תיחיוים איני	<sup>1</sup> አርርጥጥር አርር!	TTTGACGAGG(	CGCTGAACAGT
4151	GTACGAGCCC	ACAAATTGTA	TGGAAGTCG	AAACTGCTCC	CGACTTGTCA
					•
5'		11	21	31	41
1201	CAAAGTCTGG	CCTTGGGGAG	CGGTGGCTG.	L'GLLIGIGCI L'GLLIGIGCI L'GLLIGIGCI	CAAGTCCACCG STTCAGGTGGC
	GT-T-TCAGACC	GGAACCCCIC	. GCCACCOACC	- Canada Contraction	
5 '		61	71	81	91
1051	TGAAATCCTG.	ATTGTGAATT	'TGGACAACC	TGTCCTTCTT	CTTGGCCTTC AGAACCGGAAG
1251	ACTTTAGGAC	TAACACTTAA	ACCTGTTGG	CACAGGAAGAA	AGAACCGGAAG
51		11	21	31 .	41
_	CARCCA ACCT	ריים ארייייריאיי	CTTCCTCATT	TTGTCAAAAC	CACTGTGTGAT
1301	GTACGTTGGA	GGTTGAAGTA	CAACCAGTA	<b>AACAGTTTT</b>	TGACACACTA
	•			01	91
5'		61 	71	81 \	TAGTCTGCCT
351	GTTTTTATCA CDDDDDDTDGT'	HIATACIGCC TATATGACGG	TAAGGTGTAI	ACATCTCTAC	ATCAGACGGA
	Christino				Δ.
ا 5		11	21	31	41
401	GGCTTTCCTT	TCTTTAGCC	AATCGAATGC	TCTTGATCAT	GCCCTCAATC
	CCGAAAGGAA	AAGAAATCGG	TTAGCTTACG	AGAACIAGIA	CGGGAGTTAG
51	. 6	51	71		91
	rcatctctag(	TTTTATCAC	GTCTCTGCTA	ATTCCTGAAA	CTTGAATCGA GAACTTAGCT
451	AGTAGAGATC	BAAAATAGTG	CAGAGACGAT	TAAGGACTTT	GAACTTAGCT
	4	٠,	דכ	31	41
5',			21 ~aarggrgar		
501	LCV V V V CV V CV V CV V CV V CV V CV V	CCAAGTAGA	GTTACCACTA	CAAGTCAAGG	TTCTGAATCT AAGACTTAGA

5'	61	71	81	91	
4551 CATA	rcagtttctcgta Agtcaaagagcat	CTCCTCCAT GAGGAGGTA	GTCAAAGTCAC CAGTTTCAGTG	91 TGACACACTCATC ACTGTGTGAGTAG	:(C
5 '	11	21	31	41	
4601 TCAT	TTGGTGTAGGAAA AACCACATCCTTT	GCTG CTCTT CGACGAGAA	IGGTAATCAGT ACCATTAGTCA	TCCTTTAGCCAGG AGGAAATCGGTCC	Į.
5 '	61	71	81	91	
4651 CTAA	CAAAACAAGTGT	GACAGATGG(	GACTTGGTGT.	TACCTGGAAAACT ATGGACCTTTTGA	C
5 '	<b>11</b> .	21	31	41	
4701 TGTG ACAC	CTCTATTTTCTT GAGATAAAAGAA	TTCCAAAACC AAGGTTTTGC	AGGG TGTTCT TCCCACAAGA	TTTTGGGGGAAGC AAAA CCCCCTTCG	Τ
5'	61	71	81	91	
4751 ACGA	TGGGAAAGCCAA( ACCCTTTCGGTT(	SAAA GGCTAA CTTT CCGATT	AGAGAAAATG( TCTCTTTTAC(	AAATTAATGTTT CTTTAATTACAAA	C G
5'	11	21	31	41	
4801 TTTT	ACTCCCTTCAACA IGAGGGAAGTTG1	ATCAAGGTTA AGTTCCAAT	GGAA TATGTAT CCTTATACATA	TTTCATAAAAGCTI AAAGTATTTTCGA!	A P
5 '	61		81	91	
				TGGCAGGCGACTG( ACCGTCCGCTGAC(	
5'	11	21	31	41	
4901 TTGGG AACCG	GGGAG CTGGAGA CCCTCGACCTCT	GCCTTCTCT CGGAAGAGAI	TTCTTTCATGT AAGAAAGTACA	41 TGT CGTAAAAAA ACAGCATTTTTT	
5 '		71	81	. <b>91</b> .	
4951 AACGI	GAATA TGGGGCT! CTTAT ACCCCGA!	GGAAGATAA( CCTTCTATT(	CAAC TTTAACT STTG AAATTGA	CTCTTCACAGCCT GAGAAGTGTCGGA	
5 '	11	21	31	41	
5001 GCACT CGTGA	GATTTTTTCTGG/ CTAAAAAAAGACCT	ACA AATTCTT IGT TTAAGAA	CAATGGCATC GTTACCGTAG	TATTATCGCTTTT ATAATAGCGAAAA	) L
5 '	61	71	81	91	
5051 GCTAC	TACGTTTGGGTCC ATGCAAACCCAGG	ETGTTGAGCA SACAACTCGT	TTT CCTTCAA AAA GGAAGTT	91 AAA CAAAAAAAGC TTTGTTTTTCG	
5'	11	21	31	41	
5101 ACATT	rttaaaaagtcaa Aaatttttcagtt	GG TTAAGAT CCAATTCTA	CCA CCTGCAA <i>I</i> GGT GGACGTTT	AAAAAAGCTGCAA FTTTTTCGACGTT	
5 '	61	71	. 81	91	
5151 TATAAC	CGAGGAATTCTA	GTTGTCACA(	GGA AATAAAA	TGTCTGTTCCCA	

5 ¹	11	21	31	41	
5201	TATAATCAATGI LATATTAGTTAGI	TAGACTGA TAI ATCTGACTATT	ATATTATGCCI PATAATACGG	AGCAAATAGT CCGTTTATCA	TTTGAAGT AAACTTCA
5'	61	71	81	91	
5251 G	CCTAGGCACAGTG GGATCCGTGTCAC	GGAGGAGGTT CCTCCTC CAA	TTGTTCCAC( LAACAAGGTG(	SCTGTTCATA CGACAAGTAT	AGCCAATA TCGGTTAT
5 '	11	21		41	
5301 G	CCCAGCAAA AGA GGGTCGTTT TCT	CCTTAAAGGA GGAATTTCCT	CAACTTGTAA GTTGAACATT	TTTGGGACA: 'AAACCCTGT!	TTCACATC AGTGTAG
5 1	61	71	81	91	
5351 <sup>T</sup>	GTCCTCTTCATC CAGGAGAAGTAG	TGATCTGGCT ACTAGACCGA	CCCAGTGTCA GGGTCACAGT	CTCTCTAACA GAGAGATTGI	ACGGTCCT CGCCAGGA
5 '	11	21	. 31	41	
5401 T	AGAGGGACA ATT TCTCCCTGT TAA	TATCCCTGCC ATAGGGA CGG	TCTGCTTGAT AGACGAACTA	CTTATGCATO GAATACGTAO	TATCTGT 'ATAGACA
5 ' .	61		81		
5451 A	TTCTTCCAGCCA: AAGAAGGTCGGT!	rcctggcga Agggaccgct	CCTGATTTTT GGACTAAAAA	CTAAGGCACC GATTCCGTGG	CAAAACT GTTTTGA
5'	11	21	31	41	
5501 G	TAAGCTACTTCTT ATTCGATGAAGAA	TATAATCTATA TATTAGATAT	AATTCTGAGC TTAAGACTCG	ATATTAGTTA PATAATCAAT	GCCTGAG CGGACTC
5 '	. <b>61</b>	71	. 81	91	
5551 CC	CTCCAGGATATCT SAGGTCCTATAGA	TTCTTCCCTA AAGAAGGGAI	TACTCAGTC CATGAGT CAG	CAGTTTTAGC STCAAAA TCG	TGCCCAG ACGGGTC
5'	11	21	31	41	
5601 AA	AGGATTCAAAGCT CCTAAGTTTCGA	GATCTA CGAG CTAGATGCTC	TAGATCACTO ATCTAGTGAO	CCTGTCTACA GACAGATGT	GCTTGTT CGAACAA
5 '	61	71	81	91	
	'AGATCTTGTTTC' 'TCTAGAACAAAG	TCAAGCCCTG AGTTCGGGAC	GAAGCCATCA CTTCGGTAGI	GCCAGGTAA CGGTCCATT	GATTGTA CTAACAT
5'	11 -	21	31	41	
5701 AA TT	ACAATCCCTTTC: TGTTAGGGAAAGA	TAATCATGGG ATTAGTACCC	ምርምርር <i>ር ርር</i> እ	አርጥር አአጥርር	CCGGAAT EGCCTTA
5 '	61	71	81	91	
5751 TC					



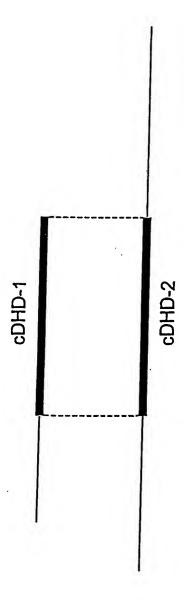


Figure 14

## Figure 15

1	CGCCCGGGCA	<b>GGTCTGTTGG</b>	AGGGCAGTTG	GTCAACCTGA	CCAGAGAGAG	CTGAGCTGGA
	GCGGGCCCGT	CCAGACAACC	TCCCGTCAAC	CAGTTGGACT	GGTCTCTCTC	GACTCGACCT
61	AGACCCCACT	GATGGTGTGC	TGCCTTTCAG	TCCAGGAAGA	AAGAAAGGAA	GGATTCTGAG
				AGGTCCTTCT		
121				TCTGTATACT		
				AGACATATGA		
181				· CGCGCCCTGA		
101				GCGCGGGACT		
241				TGGTTTGACG	~	
, , , , , , , , , , , , , , , , , , ,				ACCAAACTGC		
301				ATTTGTTTCT		
301				TAAACAAAGA		
				CAAAGCAAAA		
361				GTTTCGTTTT		
421				GGGAGTCGTG		
				CCCTCAGCAC		
481	AGAGCAGCGC (					
	TCTCGTCGCG (					
541	CATCAGGATA (					
	GTAGTCCTAT (					
601	TAGCCTGTGT (					
	ATCGGACACA C	CACAAGTATG	GTGGGCCCTA	CTTCCTTCCG	GTTGGGGCCG	AGTAGGGACG
661	AGGGCCCATC A	ACCCAGGGTA	CCACCATCTC	TGCCTACGTG	GCCAAGTCTA	GGAAGACGTT
			_			
	TCCCGGGTAG . T	rgggtcccat	GGTGGTAGAG	ACGGATGCAC	CGGTTCAGAT	CCTTCTGCAA
<b></b>	TCCCGGGTAG - 1	EcoRV	GGTGGTAGAG	) Indi	•	CCTTCTGCAA
•		EcoftV		Tho!	w	
721	, . <del></del>	EORY SATATCCTTG	GGGATGAGCG	ATTTCCTCGA	GGTACTGGCC	TGGAATCAGG
721	GTTGGTAGAG C	EORY SATATCCTTG STATAGGAAC	GGGATGAGCG CCCTACTCGC	ATTTCCTCGA TAAAGGAGCT	GGTACTGGCC CCATGACCGG	TGGAATCAGG ACCTTAGTCC
•	GTTGGTAGAG G CAACCATCTC C AACCCGCATC C	EORY EATATCCTTG CTATAGGAAC CAGTCTGTTC	GGGATGAGCG CCCTACTCGC TTTGCTTGCC	ATTTCCTCGA TAAAGGAGCT CATTGTCACT	GGTACTGGCC CCATGACCGG GCCATTGGAG	TGGAATCAGG ACCTTAGTCC ACTTGATTGG
721	GTTGGTAGAG G CAACCATCTC C AACCCGCATC C TTGGGCGTAG G	EORY SATATCCTTG TATAGGAAC CAGTCTGTTC STCAGACAAG	GGGATGAGCG CCCTACTCGC TTTGCTTGCC AAACGAACGG	ATTTCCTCGA TAAAGGAGCT CATTGTCACT GTAACAGTGA	GGTACTGGCC CCATGACCGG GCCATTGGAG CGGTAACCTC	TGGAATCAGG ACCTTAGTCC ACTTGATTGG TGAACTAACC
721	GTTGGTAGAG G CAACCATCTC C AACCCGCATC C TTGGGCGTAG G CATCCTTGAA C	EMRY EATATCCTTG CTATAGGAAC CAGTCTGTTC CTCAGACAAG TTGTACAGGC	GGGATGAGCG CCCTACTCGC TTTGCTTGCC AAACGAACGG ACTGGGGCAA	ATTTCCTCGA TAAAGGAGCT CATTGTCACT GTAACAGTGA AGAGGCCTTC	GGTACTGGCC CCATGACCGG GCCATTGGAG CGGTAACCTC TGCCTCAGCC	TGGAATCAGG ACCTTAGTCC ACTTGATTGG TGAACTAACC ATCAGGAGGT
721 781 841	GTTGGTAGAG G CAACCATCTC C AACCCGCATC C TTGGGCGTAG G CATCCTTGAA C GTAGGAACTT G	EMRY EATATCCTTG TATAGGAAC AGTCTGTTC TCAGACAAG TGTACAGGC ACATGTCGG	GGGATGAGCG CCCTACTCGC TTTGCTTGCC AAACGAACGG ACTGGGGCAA TGACCCCGTT	ATTTCCTCGA TARAGGAGCT CATTGTCACT GTARCAGTGA AGAGGCCTTC TCTCCGGAAG	GGTACTGGCC CCATGACCGG GCCATTGGAG CGGTAACCTC TGCCTCAGCC ACGGAGTCGG	TGGAATCAGG ACCTTAGTCC ACTTGATTGG TGAACTAACC ATCAGGAGGT TAGTCCTCCA
721	GTTGGTAGAG G CAACCATCTC C AACCCGCATC C TTGGGCGTAG G CATCCTTGAA C GTAGGAACTT G TGCAACAGCC A	EMRY EATATCCTTG TATAGGAAC CAGTCTGTTC ETCAGACAAG CTGTACAGGC EACATGTCGG ATCTTGCTT	GGGATGAGCG CCCTACTCGC TTTGCTTGCC AAACGAACGG ACTGGGGCAA TGACCCCGTT GGGCTTCCGT	ATTTCCTCGA TAAAGGAGCT CATTGTCACT GTAACAGTGA AGAGGCCTTC TCTCCGGAAG AGCAATACAC	GGTACTGGCC CCATGACCGG GCCATTGGAG CGGTAACCTC TGCCTCAGCC ACGGAGTCGG CAGGTGCAGG	TGGAATCAGG ACCTTAGTCC ACTTGATTGG TGAACTAACC ATCAGGAGGT TAGTCCTCCA TGTGTAGAGG
721 781 841	GTTGGTAGAG G CAACCATCTC C AACCCGCATC C TTGGGCGTAG G CATCCTTGAA C GTAGGAACTT G TGCAACAGCC A ACGTTGTCGG T	EMRY EATATCCTTG CTATAGGAAC CAGTCTGTTC CTCAGACAAG CTGTACAGGC CACATGTCGG ATCTTGCTT CTAGAACGAA	GGGATGAGCG CCCTACTCGC TTTGCTTGCC AAACGAACGG ACTGGGGCAA TGACCCCGTT GGGCTTCCGT CCCGAAGGCA	ATTTCCTCGA TAAAGGAGCT CATTGTCACT GTAACAGTGA AGAGGCCTTC TCTCCGGAAG AGCAATACAC TCGTTATGTG	GGTACTGGCC CCATGACCGG GCCATTGGAG CGGTAACCTC TGCCTCAGCC ACGGAGTCGG CAGGTGCAGG GTCCACGTCC	TGGAATCAGG ACCTTAGTCC ACTTGATTGG TGAACTAACC ATCAGGAGGT TAGTCCTCCA TGTGTAGAGG ACACATCTCC
721 781 841	GTTGGTAGAG G CAACCATCTC C AACCCGCATC C TTGGGCGTAG G CATCCTTGAA C GTAGGAACTT G TGCAACAGCC A ACGTTGTCGG T	EMNY EATATCCTTG CTATAGGAAC CAGTCTGTTC CTCAGACAAG CTGTACAGGC IACATGTCGG ATCTTGCTT TAGAACGAA AGACCGAAC	GGGATGAGCG CCCTACTCGC TTTGCTTGCC AAACGAACGG ACTGGGGCAA TGACCCCGTT GGGCTTCCGT CCCGAAGGCA	ATTTCCTCGA TAAAGGAGCT CATTGTCACT GTAACAGTGA AGAGGCCTTC TCTCCGGAAG AGCAATACAC TCGTTATGTG CCTACTCGAC	GGTACTGGCC CCATGACCGG GCCATTGGAG CGGTAACCTC TGCCTCAGCC ACGGAGTCGG CAGGTGCAGG GTCCACGTCC GTATCAAAGA	TGGAATCAGG ACCTTAGTCC ACTTGATTGG TGAACTAACC ATCAGGAGGT TAGTCCTCCA TGTGTAGAGG ACACATCTCC. CATACTTTGA
721 781 841 901 961	GTTGGTAGAG G CAACCATCTC C AACCCGCATC C TTGGGCGTAG G CATCCTTGAA C GTAGGAACTT G TGCAACAGCC A ACGTTGTCGG T TCTCGCCAAA C AGAGCGGTTT G	EMRY EATATCCTTG CTATAGGAAC CAGTCTGTTC ETCAGACAAG CTGTACAGGC AACTTGCTT TAGAACGAA AGACCGAAC TCTGGCTTG	GGGATGAGCG CCCTACTCGC TTTGCTTGCC AAACGAACGG ACTGGGGCAA TGACCCCGTT GGGCTTCCGT CCCGAAGGCA TGAATGACTT	ATTTCCTCGA TAAAGGAGCT CATTGTCACT GTAACAGTGA AGAGGCCTTC TCTCCGGAAG AGCAATACAC TCGTTATGTG CCTACTCGAC GGATGAGCTG	GGTACTGGCC CCATGACCGG GCCATTGGAG CGGTAACCTC TGCCTCAGCC ACGGAGTCGG CAGGTGCAGG GTCCACGTCC GTATCAAAGA CATAGTTTCT	TGGAATCAGG ACCTTAGTCC ACCTGATTGG TGAACTAACC ATCAGGAGGT TAGTCCTCCA TGTGTAGAGG ACACATCTCC. CATACTTTGA GTATGAAACT
721 781 841	GTTGGTAGAG CAACCATCTC AACCCGCATC TTGGGCGTAG CATCCTTGAA GTAGGAACTT GCAACAGCC ACGTTGTCGG TCTCGCCAAA CAGAGCGGTTT GTAGACAGTT TCTCGCCAAA CAGAGCGGTTT TAACATAGTT GTAACATAGTT GTAACATAGTT GCAACATAGTT TAACATAGTT	EMRY  SATATCCTTG  TATAGGAAC  AGTCTGTTC  STCAGACAAG  ACATGTCGG  ATCTTGCTT  TAGAACGAA  AGACCGAAC  TCTGGCTTG  CCATAGACT	GGGATGAGCG CCCTACTCGC TTTGCTTGCC AAACGAACGG ACTGGGGCAA TGACCCCGTT GGGCTTCCGT CCCGAAGGCA IGAATGACTT ACTTACTGAA	ATTTCCTCGA TAAAGGAGCT CATTGTCACT GTAACAGTGA AGAGGCCTTC TCTCCGGAAG AGCAATACAC TCGTTATGTG CCTACTCGAC GGATGAGCTG ACACATCATG	GGTACTGGCC CCATGACCGG GCCATTGGAG CGGTAACCTC TGCCTCAGCC ACGGAGTCGG CAGGTGCAGG GTCCACGTCC GTATCAAAGA CATAGTTTCT ATATATGCAA	TGGAATCAGG ACCTTAGTCC ACCTGATTGG TGAACTAACC ATCAGGAGGT TAGTCCTCCA TGTGTAGAGG ACACATCTCC CATACTTTGA GTATGAAACT AAAATCTAGT
721  781 841  901 961	GTTGGTAGAG CAACCATCTC AACCCGCATC TTGGGCGTAG CATCCTTGAA CGTAGGAACTT GCAACAGCC ACGTTGTCGG TCTCGCCAAA CAGAGCGGTTT GAACATAGTT GAACATAGTT GAACATAGTT GAACATAGTT CATCGTATCAA CATTGTATCAA	EMRY EATATCCTTG CTATAGGAAC AGTCTGTTC STCAGACAAG CTGTACAGGC ACATGTCCG ATCTTGCTT TAGAACGAA AGACCGAAC TCTGGCTTG CCATAGACT	GGGATGAGCG CCCTACTCGC TTTGCTTGCC AAACGAACGG ACTGGGGCAA TGACCCCGTT GGGCTTCCGT CCCGAAGGCA TGAATGACTT ACTTACTGAA GAGATGAACT	ATTTCCTCGA TARAGGAGCT CATTGTCACT GTAACAGTGA AGAGGCCTTC TCTCCGGAAG AGCAATACAC TCGTTATGTG CCTACTCGAC GGATGAGCTG ACACATCATG TGTGTAGTAC	GGTACTGGCC CCATGACCGG GCCATTGGAG CGGTAACCTC TGCCTCAGCC ACGGAGTCGG CAGGTGCAGG GTCCACGTCC GTATCAAAGA CATAGTTTCT ATATATGCAA TATATACGTT	TGGAATCAGG ACCTTAGTCC ACTTGATTGG TGAACTAACC ATCAGGAGGT TAGTCCTCCA TGTGTAGAGG ACACATCTCC CATACTTTGA GTATGAAACT AAAATCTAGT TTTTAGATCA
721 	GTTGGTAGAG CAACCATCTC AACCCGCATC TTGGGCGTAG CATCCTTGAA CGTAGGAACTT GCAACAGCC ACGTTGTCGG TCTCGCCAAA CAGAGCGGTTT CTAACATAGTT CAACATAGTT CAACATAGTT CAACATAGTT CAACCCCGAC CAACCCCCGAC CCAACCCCCGAC CAACCCCCGAC CCAACCCCCGAC CAACCCCCGAC CCAACCCCCGAC CCAACCCCCGAC CAACCCCCGAC CCAACCCCCGAC CCAACCCCCCAC CCAACCCCCGAC CCAACCCCCGAC CCAACCCCCGAC CCAACCCCCGAC CCAACCCCCCAC CCAACCCCCCAC CCAACCCCCCAC CCAACCCCCC	EMRY  EATATCCTTG  CTATAGGAAC  AGTCTGTTC  ETCAGACAAG  CTGTACAGGC  ACATGTCGG  ATCTTGCTT  TAGAACGAA  AGACCGAAC  CCATAGACT  GGTATCTGA	GGGATGAGCG CCCTACTCGC TTTGCTTGCC AAACGAACGG ACTGGGGCAA TGACCCCGTT GGGCTTCCGT CCCGAAGGCA TGAATGACTT ACTTACTGAA CTCTACTTGA GAGATGAACT TCCTCCAGGT	ATTTCCTCGA TARAGGAGCT CATTGTCACT GTAACAGTGA AGAGGCCTTC TCTCCGGAAG AGCAATACAC TCGTTATGTG CCTACTCGAC GGATGAGCTG ACACATCATG TGTGTAGTAC GGACCACAAG	GGTACTGGCC CCATGACCGG GCCATTGGAG CGGTAACCTC TGCCTCAGCC ACGGAGTCGG CAGGTGCAGG GTCCACGTCC GTATCAAAGA CATAGTTTCT ATATATGCAA TATATACGTT AACAAGGAGC	TGGAATCAGG ACCTTAGTCC ACCTGATTGG TGAACTAACC ATCAGGAGGT TAGTCCTCCA TGTGTAGAGG ACACATCTCC CATACTTTGA GTATGAAACT AAAATCTAGT TTTTAGATCA
721 	GTTGGTAGAG CAACCATCTC AACCCGCATC TTGGGCGTAG GTAGGAACTT GTAGGAACTT TGCAACAGCC ACGTTGTCGG TCTCGCCAAA CAGGCGGTTT GTAACATAGTT GAACATAGTT GAACGCCGAC CTTGCGGCTG CTTGCGGCTG CTTGCGGCTG CTTGCGGCTG CCTTGCGGCTG CCAACCCCAC CCTTGCGGCTG CCAACCCCAC CCTTGCGGCTG CCAACCCCAC CCACCCCAC CCACCCCCAC CCACCCCAC CCACCCCAC CCACCCCAC CCACCCCCAC CCACCCCCAC CCACCCCCAC CCCCCC	EMRY EATATCCTTG CTATAGGAAC CAGTCTGTTC CTCAGACAAG CTGTACAGGC ACATGTCGG ATCTTGCTT TAGAACGAA AGACCGAAC TCTGGCTTG CCATAGACT GGTATCTGA GCTGCGCGC CGACGCGCG	GGGATGAGCG CCCTACTCGC TTTGCTTGCC AAACGAACGG ACTGGGGCAA TGACCCCGTT CCCGAAGGCA TGAATGACTT ACTTACTGAA CTCTACTTGA GAGATGAACT TCTTCCAGGT AGAAGGTCCA	ATTTCCTCGA TAAAGGAGCT CATTGTCACT GTAACAGTGA AGAGGCCTTC TCTCCGGAAG AGCAATACAC TCGTTATGTG CCTACTCGAC GGATGAGCTG ACACATCATG TGTGTAGTAC TGTGTAGTAC CGGACCACAAG CCTGGTGTTC	GGTACTGGCC CCATGACCGG GCCATTGGAG CGGTAACCTC TGCCTCAGCC ACGGAGTCGG CAGGTGCAGG GTCCACGTCC GTATCAAAGA CATAGTTTCT ATATATGCAA TATATACGTT AACAAGGAGC TTGTTCCTCG	TGGAATCAGG ACCTTAGTCC ACTTGATTGG TGAACTAACC ATCAGGAGGT TAGTCCTCCA TGTGTAGAGG ACACATCTCC CATACTTTGA GTATGAAACT TATTAGATCA TGTACTCGGA ACATGAGCCT
721 781 841 901 961 1021 1081	GTTGGTAGAG CAACCATCTC AACCCGCATC CTTGGGCGTAG GTAGGAACTT GGAACAGCC ACGTTGTCGG TCTCGCCAAA CAGGCGGTTT GAACATAGTT GAACATAGTT GAACATCAT CTTGCGCGAC CTTGCGGCTG CTTGCGGCTG CCTTGCGGCTG CCTGCGGCTG CCTGCTTTGAC ACCCCATCAC CCTGCTGCTAC CCTGCTTTGAC CCAACCCCAC CCTGCCGCTG CCTGCCGCTAC CCTGCCGCTG CCTGCCGCTAC CCTGCCGCTG CCTGCCGCTAC CCTGCCGCTAC CCTGCCGCTAC CCTGCCGCTAC CCTGCCGCTAC CCTGCCGCTAC CCTGCCGCTAC CCTGCTGCGCTAC CCTGCCGCTAC CCTGCTGCCGCTAC CCTGCTGCCAC CCTGCCGCTAC CCTGCCGCTAC CCAACCCCAC CCTGCCGCTAC CCTGCTGCCAC CCTGCTGCCAC CCTGCTGCCAC CCAACCCCACAC CCAACCCCACAC CCTGCTGCCAC CCAACCCCACAC CCAACCCCACAC CCACACCCACAC CCACACCAC	EMRY EATATCCTTG CTATAGGAAC CAGTCTGTTC CTCAGACAAG CTGTACAGGC ACATGTCGG ATCTTGCTT TAGAACGAA AGACCGAAC TCTGGCTTG CCATAGACT GGTATCTGA GCTGCGCGC CGAÇGCGCG TTGGGGAGG	GGGATGAGCG CCCTACTCGC TTTGCTTGCC AAACGAACGG ACTGGGGCAA TGACCCCGTT GGGCTTCCGT CCCGAAGGCA TGAATGACTT ACTTACTGAA CCTTACTTGA GAGATGAACT TCTTCCAGGT AGAAGGTCCA AGAAGGAGGG	ATTTCCTCGA TAAAGGAGCT CATTGTCACT GTAACAGTGA AGAGGCCTTC TCTCCGGAAG AGCAATACAC TCGTTATGTG CCTACTCGAC GGATGAGCTG TGTGTAGTAC TGTGTAGTAC GGACCACAAG CCTGGTGTTC GAAGCCCATC	GGTACTGGCC CCATGACCGG GCCATTGGAG CGGTAACCTC TGCCTCAGCC ACGGAGTCGG CAGGTGCAGG GTCCACGTCC GTATCAAAGA CATAGTTTCT ATATATGCAA TATATACGTT AACAAGGAGC TTGTTCCTCG	TGGAATCAGG ACCTTAGTCC ACTTGATTGG TGAACTAACC ATCAGGAGGT TAGTCCTCCA TGTGTAGAGG ACACATCTCC CATACTTTGA GTATGAAACT TATTAGATCA TGTACTCGA ACATCTCGA ACATCTCGA CATACTCGA
721 781 841 901 961 1021 1081	GTTGGTAGAG CAACCATCTC AACCCGCATC CTTGGGCGTAG GTAGGAACTT GTAGGAACTT TCTCGCCAAA CAGGGGGTTT GTAACATAGTT GAACAGCCGAC CTTGCGGCTAC CTTGCGCCAC CTTGCGGCTG GACGCCTTGCAC CCTTGCGCCTG CCTGTTTGAC GGACACACCG GGACACACCG CTGGGACACCG CCTGCTTTGAC GGACACACCG CCAACCCCCCCCCC	EMNY EATATCCTTG CTATAGGAAC CAGTCTGTTC CTCAGACAAG CTGTACAGGC AACATGTCGG ATCTTGCTT TAGAACGAAC TCTGGCTTG CCATAGACT GGTATCTGA GCTGCGCGC CGACGCGCG AACCCCTCC	GGGATGAGCG CCCTACTCGC TTTGCTTGCC AAACGAACGG ACTGGGGCAA TGACCCCGTT GGGCTTCCGT CCCGAAGGCA TGAATGACTT ACTTACTGAA CTCTACTTGA GAGATGAACT TCTTCCAGGT AGAAGGTCCA AGAAGGTCCA AGAAGGAGGG	ATTTCCTCGA TAAAGGAGCT CATTGTCACT GTAACAGTGA AGAGGCCTTC TCTCCGGAAG AGCAATACAC TCGTTATGTG CCTACTCGAC GGATGAGCTG TGTGTAGTAC GGACCACATG TGTGTAGTAC CCTGGTGTTC GGAGCCCATC CTTCGGGTAG	GGTACTGGCC CCATGACCGG GCCATTGGAG CGGTAACCTC TGCCTCAGCC ACGGAGTCGG CAGGTGCAGG GTCCACGTCC GTATCAAAGA CATAGTTTCT ATATATGCAA TATATACGTT AACAAGGAGC TTGCAAGAAGA AAGTTCTTCT	TGGAATCAGG ACCTTAGTCC ACTTGATTGG TGAACTAACC ATCAGGAGGT TAGTCCTCCA TGTGTAGAGG ACACATCTCC CATACTTTGA GTATGAAACT AAAATCTAGT TTTTAGATCA TGTACTCGGA ACATGAGCCT CCAAGGAGAT GGTTCCTCTA
721 781 841 901 961 1021 1081	GTTGGTAGAG CAACCATCTC AACCCGCATC CTTGGGCGTAG GTAGGAACTT GTAGGAACTT TCTCGCCAAA CAGAGCGGTTT AACATAGTT GAACGCGAC CTTGCGGCTA CTTGCGCAC CTTGCGGCTG CTTGCGGCTG CTTGCGGCTG CTTGCGGCTG CTTGCGGCTG CTTTTGAC CTTGCGCTAAC CTTGCGGCTG CTTGCGGCTG CTTGCGGCTG CTTGCGGCTG CTGCTTTTGAC CAGATTTTCC CAGATTTTCC CAACCCAAC CAACCCCAC CTTGCGGCTG CTTGCGGCTG CCTGTTTGAC CAGATTTTCC CAGATTTTCC CAACCCCAC CAACCCCAC CTGCGGCTG CCTGCTTTGAC CAGATTTTCC CAGATTTTCC CAACCCCAC CCAACCCC CTGCGGCTG CCACACCCC CTGCGGCTG CCACACCCC CCTGCTTTGAC CAGATTTTCC CAGATTTTCC CAACCCCACCC CAACCCCCACCC CAACCCCCACCC CAACCCCCACCC CAACCCCCC	EMNY  SATATCCTTG  CTATAGGAAC  CAGTCTGTTC  CTCAGACAAG  CTGTACAGGC  ACCTTGCTT  TAGAACGAAC  TCTGGCTTG  CCATAGACT  GCTATCTGA  GCTGCGCGC  CTGGGGAGG  ACCCCTCC  ITGAGAAAG  ACCCCTCC  ITGAGAAAG  CTGAGAAAG  CTGAGAAAG  CTGGGAAG  CTGGGGAGG  CTGGGGAGG  CTGGGGAAG  CTGGGGAAG  CTGGGAAAAG  CTGAGAAAG  CTTGAGAAAAG  CTTGAGAAAAG  CTTGAGAAAAG  CTTGAGAAAAG	GGGATGAGCG CCCTACTCGC TTTGCTTGCC AAACGAACGG ACTGGGGCAA TGACCCCGTT GGGCTTCCGT CCCGAAGGCA TGAATGACTT ACTTACTGAA CTCTACTTGA GAGATGAACT TCTTCCAGGT AGAAGGTCCA AGAAGGTCCA GGAAGGAGGG ICTTCCTCCC GGATTGCTGG	ATTTCCTCGA TAAAGGAGCT CATTGTCACT GTAACAGTGA AGAGGCCTTC TCTCCGGAAG AGCAATACAC TCGTTATGTG CCTACTCGAC GGATGAGCTG TGTGTAGTAC GGACCACAAG CCTGGTGTTC GAAGCCCATC CTTCGGGTAG TCAAGTGGCA	GGTACTGGCC CCATGACCGG GCCATTGGAG GCCATTGGAG CGGTAACCTC ACGGAGTCGG CAGGTGCAGG GTATCAAAGA CATAGTTTCT ATATATGCAA TATATACGTT AACAAGGAGC TTGTTCCTCG TTCAAGAAGA AAGTTCTTCT AGAACAGGCG	TGGAATCAGG ACCTTAGTCC ACTTGATTGG TGAACTAACC ATCAGGAGGT TAGTCCTCCA TGTGTAGAGG ACACATCTCC CATACTTTGA GTATGAAACT AAAATCTAGT TTTTAGATCA TGTACTCGGA ACATGAGCCT CCAAGGAGAT GGTTCCTCTA AAGTCTTGAA
721 	GTTGGTAGAG CAACCATCTC AACCCGCATC CTTGGGCGTAG GTAGGAACTT GTAGGAACTT TCCAACAGCC ACGTTGTCGG TACATAGTT GAACATAGTT GAACGCCGAC CTTGCGGCTG CCTGTTTGAC ATGTATCAA CCTTGCGGCTG CCTGTTTGAC CAGATTTTCC ATGGACAAACTG CAGATTTTCC ATGTTAAAAGG TI	EMNY  SATATCCTTG  CTATAGGAAC  CAGTCTGTTC  STCAGACAAG  CTGTACAGGC  ACCTTGCTT  TAGAACGAA  AGACCGAAC  TCTGGCTTG  CCATAGACT  GGTATCTGA  GCTGCGCGC  CGACGCGCG  TTGGGGAGG  AACCCCTCC  TTGAGAAAG  AACTCTTTC	GGGATGAGCG CCCTACTCGC TTTGCTTGCC AAACGAACGG ACTGGGGCAA TGACCCCGTT CCCGAAGGCA TGAATGACTT ACTTACTGAA CTCTACTTGA GAGATGAACT TCTTCCAGGT AGAAGGTCCA AGAAGGTCCA AGAAGGTCCA CCTTCCTCCC GGATTGCTGG CCTAACGACC	ATTTCCTCGA TAAAGGAGCT CATTGTCACT GTAACAGTGA AGAGGCCTTC TCTCCGGAAG AGCAATACAC TCGTTATGTG CCTACTCGAC GGATCACTG ACACATCATG TGTGTAGTAC CCTGGTGTTC CCTGGTGTTC CAAGCCCATC CTTCGGGTAG TCAAGTGGCA AGTTCACCGT	GGTACTGGCC CCATGACCGG GCCATTGGAG GCGTAACCTC TGCCTCAGCC ACGGAGTCGG CAGGTGCAGG GTCCACGTCC GTATCAAAGA CATAGTTTCT ATATATGCAA TATATACGTT AACAAGGAGC TTCAAGAAGA AAGTTCTTCT AGAACAGGCG TCTTGTCCGC	TGGAATCAGG ACCTTAGTCC ACTTGATTGG TGAACTAACC ATCAGGAGGT TAGTCCTCCA TGTGTAGAGG ACACATCTCC CATACTTTGA GTATGAAACT TTTTAGATCA TGTACTCGGA ACATGAGCCT CCAAGGAGAT GGTTCCTCTA AAGTCTTGAA TTCAGAACTT
721 	GTTGGTAGAG CAACCATCTC AACCGGCATC CTTGGGCGTAG GTAGGAACTT GTGCAACAGCC ACGTTGTCGG TCTCGCCAAA CAGAGCGGTTT GAACATAGTT GAACGCCGAC CCTGCTGCGCTG GGACAAACTG CCTGTTTGAC CAGATTTTCC CAGATTTTCC ACGTTTTCC CAGATTTTCC CAGATTTTCC CAGATTTTCC CAGATTTTCC CATTCCCGAT CCATTCCCGAT CCATTCCCGAT CCATTCCCGAT CCATTCCCGAT CCATTCCCCGAT CCATCCCCGAT CCATCCCCCAT CCATCCCCCCCC	EMNY  SATATCCTTG  CTATAGGAAC  CAGTCTGTTC  STCAGACAAG  CTGTACAGGC  ACCTTGCTT  TAGAACGAAC  TCTGGCTTG  CCATAGACT  GGTATCTGA  GCTGCGCGC  CTGGGGAGG  ACCCCTCC  TTGAGAAAG  AACCCCTTC  CCTACGCGG	GGGATGAGCG CCCTACTCGC TTTGCTTGCC AAACGAACGG ACTGGGGCAA TGACCCCGTT GGGCTTCCGT CCCGAAGGCA TGAATGACTT ACTTACTGAA CTCTACTTGA GAGATGAACT CTTCCAGGT AGAAGGTCCA AGAAGGTCCA CCTTCCTCCC GGATTGCTGG CCTAACGACC	ATTTCCTCGA TAAAGGAGCT CATTGTCACT GTAACAGTGA AGAGGCCTTC TCTCCGGAAG AGCAATACAC TCGTTATGTG CCTACTCGAC GGATCACTG ACACATCATG TGTGTAGTAC CCTGGTGTTC CCTGGTGTTC CAAGCCCATC CTTCGGGTAG TCAAGTGGCA AGTTCACCGT TAACAGGGAG	GGTACTGGCC CCATGACCGG GCCATTGGAG GCCATTGGAG CGGTAACCTC TGCCTCAGCC ACGGAGTCGG CAGGTGCAGG GTCCACGTCC GTATCAAAGA CATAGTTTCT ATATATGCAA TATATACGTT AACAAGGAGC TTCAAGAAGA AAGTTCTTCT AGAACAGGCG TCTTGTCCGC GTGGACCTGT	TGGAATCAGG ACCTTAGTCC ACTTGATTGG TGAACTAACC ATCAGGAGGT TAGTCCTCCA TGTGTAGAGG ACACATCTCC CATACTTTGA GTATGAAACT TTTTAGATCA TGTACTCGGA ACATGAGCCT CCAAGGAGAT GGTTCCTCTA AAGTCTTGAA TTCAGAACTT ACACAGGCTA
721 781 841 901 961 1021 1081 1141	GTTGGTAGAG CAACCATCTC AACCCGCATC TTGGGCGTAG GTAGGAACTT GTAGGAACTT TGCAACAGCC ACGTTGTCGG TTCTCGCCAAA CAGACGCGTT GAACATCAT ACGTTGTACAA CATTGTATCAA CATTGTATCAA CATTGTATCAA CATTGTATCAA CATTGTATCA GAACGCCGAC CCTTGCGGCTG CCTGTTTGAC CAGATTTTCC ATGCAAAAACTG CATTCCCGAT CATTCCCGAT CATTCCCGAT CATTCCCGAT CCTAAAAGG CTAAAGGGCTA CCATTCCCGAT CCATCCCGAT CCATCCCGAT CCATCCCGAT CCATCCCGAT CCATCCCGAT CCATCCCGAT CCATCCCGAT CCATCCCGAT CCATCCCCGAT CCATCCCCCAT CCATCCCCCCAT CCATCCCCCAT CCATCCCCCAT CCATCCCCCAT CCATCCCCCAT CCATCCCCCCAT CCATCCCCCCAT CCATCCCCCCAT CCATCCCCCCAT CCATCCCCCCAT CCATCCCCCCCC	EMRY  SATATCCTTG  CTATAGGAAC  AGTCTGTTC  STCAGACAAG  CTGTACAGGC  ACATGTCGG  ATCTTGCTT  TAGAACGAA  AGACCGAAC  CCATAGACT  GGTATCTGA  GCTGCGCGC  CTGGGGAGG  AACCCTCC  TTGAGAAAG  AACTCTTTC  CCTACGCGG  GGATGCGCG  CCTACGCGG  GGATGCGCG  GGATGCGCG  CCTACGCGG  GGATGCGCC  CCTACGCGG  GGATGCCCC  CCTACGCGC  CCTACGCGG  GGATGCCCC  CCTACGCCGC  CCTACGCCGC  CCTACGCCGC  CCTACGCCGC  CCTACGCCGC  CCTACGCCGC  CCTACGCCGC  CCTACGCCGC  CCTACGCCGC  CCTACGCCCC  CCTACGCCGC  CCTACCCCC  CCTACCCCC  CCTACCCC  CCTACCCCC  CCTACCCC  CCCTACCC  CCTACCCC  CCTACCCC  CCTACCCC  CCTACCCC  CCTACCC  CCTACC  CCTACCC  CCTACCC  CCTACCC  CCTACCC  CCTACCC  CCTACCC  CCTACCC  CCCTACCC  CCCC  CCCC  CCCC  CCCC  CCCC  CCCC  CCCC	GGGATGAGCG CCCTACTCGC TTTGCTTGCC AAACGAACGG ACTGGGGCAA TGACCCCGTT GGGCTTCCGT CCCGAAGGCA TGAATGACTT ACTTACTGAA CTCTACTGA GAGATGAACT TCTTCCAGGT AGAAGGTCCA AGAAGGAGGG CCTTCCTCCC GGATTGCTGG CCTAACGACC ACCCTCGCTT CGGGAGCGAA	ATTTCCTCGA TARAGGAGCT CATTGTCACT GTAACAGTGA AGAGGCCTTC TCTCCGGAAG AGCAATACAC TCGTTATGTG CCTACTCGAC GGATGAGCTG ACACATCATG TGTGTAGTAC CCTGGTGTTC CTTCGGGTAG CCTGGTGTTC TCAAGTGGCA AGTTCACCGT TAACAGGGAG ATTGTCCCTC	GGTACTGGCC CCATGACCGG GCCATTGGAG GCCATTGGAG CGGTAACCTC TGCCTCAGCC ACGGAGTCGG GTCCACGTCC GTATCAAAGA CATAGTTTCT ATATATGCAA TATATACGTT ACAAGAAGA TTCAAGAAGA AAGTTCTTCT AGAACAGGCG TCTTGTCCGC GTGGACCTGT CACCTGGACA	TGGAATCAGG ACCTTAGTCC ACCTTAGTCC ACTTGATTGG TGAACTAACC ATCAGGAGGT TAGTCCTCCA TGTGTAGAGG ACACATCTCC CATACTTTGA GTATGAAACT AAAATCTAGT TTTTAGATCA TGTACTCGGA ACATGAGCCT CCAAGGAGAT GGTTCCTCTA AAGTCTTGAA TTCAGAACTT ACACAGGCTA TGTGTCCGAT
721 781 841 901 961 1021 1081 1141 1201	GTTGGTAGAG CAACCATCTC AACCCGCATC TTGGGCGTAG CATCCTTGAA CGTAGGAACTT GTCAACAGCC AGAGCGGTTT GTAACATAGTT GAACATAGTT GAACGCCGAC CTTGCGGCTG CCTGTTTGAC GGACAAACTG CAGATTTTCC AT GTTAAAAGG TAACGTTAAAAGG CATTCCCGAT CATTCCCGAT CACCACGAGG CAACCACGAGG CAACCACGAGG AACCCACGAGG CAACCACCAGAGG CAACCACCAGAGG CAACCACCACGAGG AACCCACCACGAGG CAACCACCACGAGG AACCCACCACGAGG CAACCCACCACGAGG AACCCACCACGAGG CAACCCACCACGAGG AACCCACCACGAGG CAACCCACCACGAGG AACCCACCACGAGG CAACCACCACGAGG AACCCACCACGAGG AACCCACCACGAGG CAACCCACGAGG AACCCACCACGAGG CAACCCACGAGG AACCCACCACGAGG AACCCACCACCACCACCACCACCACCACCACCACCACCA	EMRY  SATATCCTTG  TATAGGAAC  AGTCTGTTC  STCAGACAAG  TGTACAGGC  ACATGTCGG  ATCTTGCTT  TAGAACGAA  AGACCGAAC  CCATAGACT  GGTATCTGA  GCTGCGCGC  ACCCCTCC  TTGAGAAAG  ACCCCTCC  CCTACGCGG  GGATGCGCG  CCTACGCGG  ACCTCTTC  CCTACGCGG  ACATTCTGT  ACATTCTCT  ACATTCTCT  ACATTCT  ACATTC	GGGATGAGCG CCCTACTCGC TTTGCTTGCC AAACGAACGG ACTGGGGCAA TGACCCCGTT GGGCTTCCGT CCCGAAGGCA TGAATGACTT ACTTACTGAA CTCTACTTGA GAGATGAACT TCTTCCAGGT AGAAGGAGGG TCTTCCTCCC GGATTGCTGG CCTAACGACC ACCCTCGCTT AGGGAGCGAA TTATGCCCAT	ATTTCCTCGA TARAGGAGCT CATTGTCACT GTAACAGTGA AGAGGCCTTC TCTCCGGAAG AGCAATACAC TCGTTATGTG CCTACTCGAC GGATGAGCTG ACACATCATG TGTGTAGTAC CCTGGTGTTC CAAGCCCATC CTTCGGGTAG TCAAGTGGCA AGTTCACCGT TAACAGGGAG ATTGTCCCTC AGTGAGCCA	GGTACTGGCC CCATGACCGG GCCATTGGAG GCCATTGGAG CGGTAACCTC TGCCTCAGCC ACGGAGTCGG GTCCACGTCC GTATCAAAGA CATAGTTTCT ATATATGCAA TATATACGTT ACAAGGAGC TTCAAGAAGA AAGTTCTTCT AGAACAGGCG TCTTGTCCGC GTGGACCTGT CACCTGGACA GGCAGCGTGA	TGGAATCAGG ACCTTAGTCC ACCTTAGTCC ACTTGATTGG TGAACTAACC ATCAGGAGGT TAGTCCTCCA TGTGTAGAGG ACACATCTCC CATACTTTGA GTATGAAACT TTTTAGATCA TGTACTCGGA ACATGAGCCT CCAAGGAGAT GGTTCCTCTA AAGTCTTGAA TTCAGAACTT ACACAGGCTA TGTGTCCGAT TGTGTCCGAT
721 781 841 901 961 1021 1081 1141 1201	GTTGGTAGAG CAACCATCTC AACCCGCATC TTGGGCGTAG GTAGGAACTT GTAGGAACTT TGCAACAGCC ACGTTGTCGG TTCTCGCCAAA CAGACGCGTT GAACATCAT ACGTTGTACAA CATTGTATCAA CATTGTATCAA CATTGTATCAA CATTGTATCAA CATTGTATCA GAACGCCGAC CCTTGCGGCTG CCTGTTTGAC CAGATTTTCC ATGCAAAAACTG CATTCCCGAT CATTCCCGAT CATTCCCGAT CATTCCCGAT CCTAAAAGG CTAAAGGGCTA CCATTCCCGAT CCATCCCGAT CCATCCCGAT CCATCCCGAT CCATCCCGAT CCATCCCGAT CCATCCCGAT CCATCCCGAT CCATCCCGAT CCATCCCCGAT CCATCCCCCAT CCATCCCCCCAT CCATCCCCCAT CCATCCCCCAT CCATCCCCCAT CCATCCCCCAT CCATCCCCCCAT CCATCCCCCCAT CCATCCCCCCAT CCATCCCCCCAT CCATCCCCCCAT CCATCCCCCCCC	EMRY  SATATCCTTG  TATAGGAAC  AGTCTGTTC  STCAGACAAG  TGTACAGGC  ACATGTCGG  ATCTTGCTT  TAGAACGAA  AGACCGAAC  CCATAGACT  GGTATCTGA  GCTGCGCGC  ACCCCTCC  TTGAGAAAG  ACCCCTCC  CCTACGCGG  GGATGCGCG  CCTACGCGG  ACCTCTTC  CCTACGCGG  ACATTCTGT  ACATTCTCT  ACATTCTCT  ACATTCT  ACATTC	GGGATGAGCG CCCTACTCGC TTTGCTTGCC AAACGAACGG ACTGGGGCAA TGACCCCGTT GGGCTTCCGT CCCGAAGGCA TGAATGACTT ACTTACTGAA CTCTACTTGA GAGATGAACT TCTTCCAGGT AGAAGGAGGG TCTTCCTCCC GGATTGCTGG CCTAACGACC ACCCTCGCTT AGGGAGCGAA TTATGCCCAT	ATTTCCTCGA TARAGGAGCT CATTGTCACT GTAACAGTGA AGAGGCCTTC TCTCCGGAAG AGCAATACAC TCGTTATGTG CCTACTCGAC GGATGAGCTG ACACATCATG TGTGTAGTAC CCTGGTGTTC CAAGCCCATC CTTCGGGTAG TCAAGTGGCA AGTTCACCGT TAACAGGGAG ATTGTCCCTC AGTGAGCCA	GGTACTGGCC CCATGACCGG GCCATTGGAG GCCATTGGAG CGGTAACCTC TGCCTCAGCC ACGGAGTCGG GTCCACGTCC GTATCAAAGA CATAGTTTCT ATATATGCAA TATATACGTT ACAAGGAGC TTCAAGAAGA AAGTTCTTCT AGAACAGGCG TCTTGTCCGC GTGGACCTGT CACCTGGACA GGCAGCGTGA	TGGAATCAGG ACCTTAGTCC ACCTTAGTCC ACTTGATTGG TGAACTAACC ATCAGGAGGT TAGTCCTCCA TGTGTAGAGG ACACATCTCC CATACTTTGA GTATGAAACT TTTTAGATCA TGTACTCGGA ACATGAGCCT CCAAGGAGAT GGTTCCTCTA AAGTCTTGAA TTCAGAACTT ACACAGGCTA TGTGTCCGAT TGTGTCCGAT

1381	GCAGATGGTG AACAAGATCA GCGGTAGCGC	CTTCTCCAAG ACAGACGAGA ACAACTTCAA
	CGTCTACCAC TTGTTCTAGT CGCCATCGCG	GAAGAGGTTC TGTCTGCTCT TGTTGAAGTT
		Phres
1443	CARCTETCOT CTCTTCTCCC CACTGGCCTT	GCACTGTGCT AACATGTACC ACAGGATCCG
1441	CONTRACT CACABGACGC GTGACCGGAA	CGTGACACGA TTGTACATGG TGTCCTAGGC
	CTACAAACGA CAGAAGACGG GIGING	Hendu
	The second secon	GGAGAAGCTT TCCTACCACA GCATCTGCAC
1501	CCACTCAGAA TGCATCTACA GGGTIACCAL	CCTCTTCGAA AGGATGGTGT CGTAGACGTG
	GGTGAGTCTT ACGTAGATGT CCCAATGGTA	Chaccators Concentration
1561	CTCCGAGGAG TGGCAAGGCC TCATGCGCTT	CARCCTACCA GCACGCATCT GCCGGGACAT
	GAGGETECTE ACCETTOOGG AGTACGUGAA	GTTGGATGGT CGTGCGTAGA CGGCCCTGTA
1621	CGAGCTATTC CACTTGACA TTGGTCCTTT	CGAGAACATG TGGCCTGGGA TCTTTGTCTA
	CONCERNAC CTCARACTCT RACCAGGARA	CCICIICING NOODOOCCI NOUVICHOUI
1681	TOTAL CONTRACTOR CONTRACTOR	TTTTGAACIT GAAAAAIIGI GCCGITITAT
	CONTROL CONTRACTOR COTTAGGAC	ARACTIGAR CITITIANCA COOCHAMITA
1741	THE PROPERTY OF THE PROPERTY O	TOOTTACCAC AACTGGAAGC ATGCAGTCAC
1.41	CTACAGACAC TECTTCTTGA TAGCOGOCCA	AGGAATGGTG TTGACCTTCG TACGTCAGTG
		ARRCARCART GGCCTCTTCA CAGACCTCGA
1801	GGTGGCACAC TGCATGTATG CONTINUES	TTTGTTGTTA CCGGAGAAGT GTCTGGAGCT
	Md — — — — — — — — — — — — — — — — — — —	CCATGACCTG GACCACAGGG GCTTCAGTAA
1861	GCGCAAAGGC CTGCTAATTG CGTGTCTGTG	GGTACTGGAC CTGGTGTCCC CGAAGTCATT
	CGCGTTTCCG GACGATTAAC GCACAGACAC	COCCOCCTG TACTCCACCT CCACCATGGA
1921	CAGCTACCTG CAGAAGTTCG ACCACCCCT	GGCGCGCTG TACTCCACCT CCACCATGGA CCGCCGCGAC ATGAGGTGGA GGTGGTACCT
	GTCGATGGAC GTCTTCAAGC TGGTGGGGGA	CONTROL COME GRACEGUAGA ATATCTTCTC
1981	GCAACACCAC TTCTCCCAGA CGGTGTCCAT	CCTTCAGCTG GAAGGGCACA ATATCTTCTC
		CAMAGICONO CARGO CARROLLE
2041		CYMYS-BISBNI: MILLIANANG COMESSION
	· · · · · · · · · · · · · · · · · · ·	TINCCICING SINGUETES
2101		
2202		
2161		
5101		
	TOTAL STORY STORY CONTINUE CONTAINED	GCCAGITACA MANITONON
2221	GACACTAGAA ACGAGACACT GGTTTGATAC	CGGTCAATGT TTTAACTGTC GCTTACTATA
	***************************************	GATGAAGAAG CTGGGCATAC AGCCCATTCC
2281		
	TATACGTCTT AAGACCCCAC 1500	CCCTCAAGGG CAGCTCGGAT TCTACAATGC GGGAGTTCCC GTCGAGCCTA AGATGTTACG
2341		
	ATACTACCTG TCTCTGTTCG CTCTACTC	GCAGATCCTC CCACCCACAG AGCCTCTGCT
2401		
	ACACCGGTAA GGGACGATAT GGTGGAACTG	CONCARGETA ATTCGCGGGG AAGAGACAGC
2461		GGAGAAGGTA ATTCGCGGGG AAGAGACAGC CCTCTTCCAT TAAGCGCCCC TTCTCTGTCG
	CTTCCGGACG TCCCTATTGG AGTTAGTCAC	TAGENTICAGE ACACCTGAGA AGCTGAACGT
2521	AATGTGGATT TCAGGCCCAG GCCCGCCCC	TAGCAAGAGC ACACCTGAGA AGCTGAACGT ATCGTTCTCG TGTGGACTCT TCGACTTGCA
	TTACACCTAA AGTCCGGGTC CGGCCGCGC	CONCENTRACE GOCCAGCAAC CGACTCAACC
2581	GAAGGTTGAA GACTGATOCT GAAGTGAOGT	CCTGATGTCT GCCCAGCAAC CGACTCAACC GGACTACAGA CGGGTCGTTG GCTGAGTTGG
	CTTCCAACTT CTGACTAGGA CTTCACTGCA	TACCORONA ANCOCCUTCE CAGAAGGTAC
2641	TECTTOTETE ACTIOETTCT TITTETTTC	AAGGGGTGAA AACCCCTGT CAGAAGGTAC TTCCCCACTT TTGGGGGACA GTCTTCCATG
	ACGAAGACAC TGAAGCAAGA AAAACAAAAG	

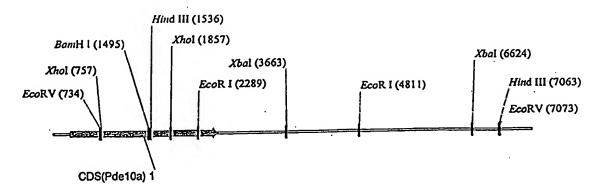
2701	CCTCCCATAT CCATGTGAAG CAGACGACTC	CCTGCTTGCC GCACACACCT CGGACAGTGA
2701	GCAGCGTATA GGTACACTTC GTCTGCTGAG	GGACGAACGG CGTGTGTGGA GCCTGTCACT
2761	CCARCCAGG CTCTGCCGTG TTCAGACGTC	GGCTACTCCG TGGCTCCACC TGACCTCCGA
2701	CCTTGGGTCC GAGACGGCAC AAGTCTGCAG	CCGATGAGGC ACCGAGGTGG ACTGGAGGCT
2821	A WOOTH TITTE CTCCCAGGCC AGCACTGCAC	TGTCTGGAGG GGGCAGAGAC CACAGGAGAG
2021	TACCATARAC GAGGGTCCGG TCGTGACGTG	ACAGACCTCC CCCGTCTCTG GTGTCCTCTC
	CERTIFICATION CONTINUES ATTENDED	GGCCAGTTCC CTAGTTCTGT GCCATGCTGC
2881	Chachacea Cetageage TacteccACA	CCGGTCAAGG GATCAAGACA CGGTACGACG
	TOTAL CANTECTER CANTEGGACA	CACGCCCTT GTTGTGAAGT TTACATGTGA
2941	ACCARCAGO CATIOGITAG GAATGOGIGG	GTGCGGGAA CAACACTTCA AATGTACACT
	ACGRACIACO GIINIGATIO GIATORICO CONTROLO CONTROL	TGGACACATG TAATGAAGGT CACAGTCCAC
3001	CCTTCTTATA GGITAACTGA GTITGTGGGG	ACCTGTGTAC ATTACTTCCA GTGTCAGGTG
	GGAAGATAT COATTOACT GRATIOTO	CAGGTGCACT ACAGGTATGC TCTTTCAGTC
3061	AGGTGACAGA GAAATCCAAA CIGIIGAIIA	GTCCACGTGA TGTCCATACG AGAAAGTCAG
	TCCACTGTCT CTTTAGGTTT GACAGGCTCC	ACTCAGAANN AAGCATACCT CTGCCCTCAT
3121	TATCTGGGGG CACATAGGTG AGTCTGCTCC	TGAGTCTTNN TTCGTATGGA GACGGGAGTA
	ATAGACCCCC GIGIATCCAC TCAGACGACC	GGGGAACTGA AGCTCTCACT TCAAACCATG
3181	CCAGGGGACA CAGGGTACAT CCCAGGCATC	CCCCTTGACT TOGAGAGTGA AGTTTGGTAC
	GGTCCCCTGT GTCCCATGTA GGGTCCGTAG	CACTGTAGCC TTCGACAACT GCGCCAATCC
3241	TCAAAGAATT AAAACACCTC CCCTCCCCCT	GTGACATCGG AAGCTGTTGA CGCGGTTAGG
	AGTTTCTTAA TTTTGTGGAG GGGAGGGGGA	ATAAATTTCC TCCAGCAAGC AAATCTTGTG
3301	CTTTATACAA AGAAAATAAA AGTAAGGCAL	TATTTAAAGG AGGTCGTTCG TTTAGAACAC
	GRANTATGIT TCTTTTATTT TCATTCCGT	CHTCTANANT NTCNCHGNAT GTTATGGCAG
3361	GGTAAAAAA AAGCATGTGA ATNITAACAA	GNAGATNINA NAGNGNCNIA CAATACCGIC
	CCATTITIT TIGHACACI TANNALIGI	ATTCCAGAAG ATACCTCATC CTATGCCTGA
3421	AATTTAGTC ACGICCAAAA CAAAAAGAII	TAAGGTCTTC TATGGAGTAG GATACGGACT
	TTAAAATCAG IGCAGGIIII GIIIICCCCA	GGGTTCTGAT CCGTCTCCTC ACGGTGCAAT
3481	AAGGCTCCAC AGCATGGCGT CCGTCTCCCA	CCCAAGACTA GGCAGAGGAG TGCCACGTTA
	TICCGAGGIG ICGIACCOCA CCTCCAGGGC	TACCACATTG ACCCAGAAGG TATCTCCTCT
3541	CAGGCAGGAC AGAGAGGAGG GCIGCAGGGC	ATGGTGTAAC TGGGTCTTCC ATAGAGGAGA
	GICCGICCIG ICICCICCIO CONTOCCON	ATGCTGTATT GAATAGTTCT CTGTGTGACT
3601	CACCATICAG ACATOCATAA GGAATGOOTE	TACGACATAA CTTATCAAGA GACACACTGA
	The state of the s	
	WWWW CONCRETE CONCRETE CONTRACTOR	TCCNGGGGAA CTCTAAGGAG TCACAGGTTC
3661	TTCTAGAGAA GCCAGGACAC CCIGAGGGI	AGGNCCCCTT GAGATTCCTC AGTGTCCAAG
	TOTAL CAMPAGE CO ATACCATGGA	CACAGAGATC CGGTCGTTGT TCTCACTCGT
3721	ACACCGTGGG GATTICAGG ATACCTTON	CTGTCTCTAG GCCAGCAACA AGAGTGAGCA
	TO THE PARTY OF TH	CACTCACTCA GCACTCTGCA GGAGCAGGAG
3781	GAGCCTTGAG AAGGAGAGAC IGACCAGAII	GTGAGTGAGT CGTGAGACGT CCTCGTCCTC
	TOTAL STORM STORMS STORMS TAGET	TTTGATACAC CCAATACCAT ACACACAGGA
3841	MAGATACITI MAGAIGMAIC IIGOMINOM	AAACTATGTG GGTTATGGTA TGTGTGTCCT
	TOTAL MICHAEL STORM ANTICACTUTE	CTTCCGCGCT CTGACCCACG GIIGIAGCGG
3901	CONCCETAN ACCTTTCAGA TAAGTCAAAG	GAAGGCGCGA GACTGGGTGC CANGITOGO
:	TOTAL	CGATTTCCCC ATGGGCTTCT AAAATGTCAC
	TO TOTAL TENED TO TENED TEN	GCTAAAGGGG TACCCGAAGA
4003	THE TOTAL CONCENTRATION OF A CONCENTRATION OF THE C	TTACTGGTTA CAAGGTGATG TCAACAAGAG
4021	GTAGAGGAGG GGACGACACA GGATGAGGTA	AATGACCAAT GTTCCACTAC AGTTGTTCTC

		•
4081		GTGCACACAC ATGTATGCAC AAGCACACAG
		CACGTGTGTG TACATACGTG TTCGTGTGTC
4141		CCCCAAAAGG AGAGAAAACAT
		GGGGTTTTCC TCTCTTTTCC TTCTTTGTA
4201		AATAGTCTTT CCTGTAGGAA ACAGGAGCTC
	AATATTTTC GCTGTCGATG GGGTATAGTT	, . ,
4261		CCATCAGTGC ACTCTCCCAG GGGTGCTCAC
		GGTAGTCACG TGAGAGGGTC CCCACGAGTG
4321		CACTGGCTGC AGCAAAAAGC CATTCGATCC
	ACTTCGACCA GGTGRAGATA TTTGTCCACT	GTGACCGACG TCGTTTTTCG GTAAGCTAGG
4381	ACACAAATTG ATCTTCTATC ATCTTGGAAT	CTGAATTGCA GGGAGGAGCA GYATGTAAGA
	TGTGTTTAAC TAGAAGATAG TAGAACCTTA	GACTTAACGT CCCTCCTCGT CYTACATTCT
4441	CGACCGTTTA ATTCAGGCAT TCCGAAGGCA	TGAGCGCATG GATTCTRTCA CCAAGCGTAT
		ACTCGCGTAC CTAAGARAGT GGTTCGCATA
 4501		CGGACTGTTT TTGCTGTAGA AGTAGGGATT
		GCCTGACAAA AACGACATCT TCATCCCTAA
4561		TGGGGCAGTT TTGCAGAGGA ACCTGCCAGA
4301	AATGTCTTCA GAGGAACRTA AACGGGACGG	ACCCCGTCAA AACGTCTCCT TGGACGGTCT
4621		TAGTATCATG TGAGAAACAG TTTGTAGAAA
4021	CTANATARC GACCAGTCAG AGAACACTTT	ATCATAGTAC ACTCTTTGTC AAACATCTTT
4681	BARACTATAC CTCCCAACAC CTTTCCAACA	TTGTTCCTTC CATGGGCCAA GACTCAGTTA
4 001	TTTTGITHTG GACCCTTCTG GAAACGTTGT	AACAAGGAAG GTACCCGGTT CTGAGTCAAT
4741		GCCAGGATAC AGCCATGTTT AGTTAATAAT
4/41	CCTCCCTATT TAGACGGCC TTATTTGATC	CGGTCCTATG TCGGTACAAA TCAATTATTA
	Ecoli Indiana	
4801	TTCCTTTTEC AATTCACACA GCCAGGATTG	GTTTTTTGT GTCTTGGCAA GTGGAGCATA
4001	AACCAAAATC TTAAGTGTGT CCGTCCTAAC	CAAAAAAACA CAGAACCGTT CACCTCGTAT
4061	TTTANCATEC ACCCATCCE ATCCTCCTC	TTAGCTTTTC CCACCCTCTT GTCTCACCAA
4861	ANATTOTATO TOCOTACOCT TAGGACGAG	AATCGAAAAG GGTGGGAGAA CAGAGTGGTT
4001	COMMONWOOD TOOR RECEPT TOORGRAPT	TCTCATTAAT GGCTGATGCA AACTTAGTGA
4921	Chanangae AGCTTTCCAR AGGTCCTTAR	AGAGTAATTA CCGACTACGT TTGAATCACT
4001	AMARAGA ATATABACAA TCCTCACCTC	ACCAAAATTA TATTATTTĞC AGTCATTTĞT
4981	TATALTACE ACCACTOGAGE	TGGTTTTAAT ATAATAAACG TCAGTAAACA
:-	CARLACIA A BROWN BOOK A BROCKTATT	ATTTAATTTG TGGCCACACA CTGTGGTTAT
5041	CTATCACTA TABBATACCE TTACCAATAA	TAAATTAAAC ACCGGTGTGT GACACCAATA
	CIAITOTOTI COMPONENCE CACALATET	TCTTGGATAT GTAAGTGCCA ATACCAGTGT
5101	CARACARCA CCARCARGA CTCTTTTACA	AGAACCTATA CATTCACGGT TATGGTCACA
	CAMACANCA COMPONENTA CARABATACAG	CCTAAGGTTT GTAAACATCA ATTCTATCTC
5161	CEMCAMARCA ACCCCCCCCC CTTTATGTC	GGATTCCAAA CATTTGTAGT TAAGATAGAG
	AGTTCATCAG AGGGCCTGAG AAGCTGCGGG	GCAGTGTAAA GTAAAGTATG CTGGGCTGGT
5221	TORROTTO TOCOGRACTO TTCGACGCCC	CGTCACATTT CATTTCATAC GACCCGACCA
	GGTGGTCAGC CTCCCCTTGC CAAGAAGAGA	GCAATTGAAT CCTGTCCCCA GCTCCCTCCA
5281	CCACCAGTCG GAGGGGAACG GTTCTTCTCT	CGTTAACTTA GGACAGGGGT CGAGGGAGGT
	CGCCTGAAGA GTGACCAGTG CTGGCCCGAC	GGATCGCTGA GATATTCTCC CATAATGGCA
5341	GCCTGAAGA GTGACCAGTG CTGGCCCGAC GCGGACTTCT CACTGGTCAC GACCGGGCTG	CCTAGCGACT CTATAAGAGG GTATTACCGT
	AAAAATAGG CAGTTTGATG TGACCTGTTT	AGTGTGGCTC TCCTCTTTTG AGCATGTGTT
5401	TTTTTTATCC GTCAAACTAC ACTGGACAAA	TCACACCGAG AGGAGAAAAC TCGTACACAA
	TITITATIC GEOMETRIC NOTONICE IN	

TCGTANAMAT AMANTATION ANGERCATION CITTOTTOTT CCARGITGIT TCATGATAGE  5521 GCTAGATATA TATAGCACAGC CIGCOTTCTG GCAGCAGAGA GGTTCACCAC AGTACATACA  5521 GCTAGATATA TATAGCACAGC CIGCOTTCTG CIGCACAACG CCTTAGAGAC CCGCCTTTC  CGATCTATATA AATCGTGTGG GACGGAAGAC GACGTGTTGC GGAATCCTG GGCCGGAAAC  5581 AATGAGCATA GCTTGTGCTC CTTTTCTGCT CTCTTAGGTCT TAAACATAGG GCCGGAAACA  5582 AATGAGCATA GCTTGCCTC CTTTTCCTGCT CTTTAGAGCCT TAAACATAGG GCCGGAAACA  5583 AATGAGACAA AAGTATGCAT CTTGCCTTGG CTTGAGCCTT TTCGTTTCTA ATGCTGATTT  TTATCTGTTT TTCATACCATA GAACGGAACC GAACTGGGAA AAGGAAAAGT TACGACTGAAA  5761 AGGGGGAAAA AGGAACCAC ACTGCATACC CTTCCCTGT TCATAACG  GAGGGGAAAA GAGGACACC ACTGCATACC TTTCCAGAGT GTAAAGGACA ACTTTAAAC  5762 AGGCGTGTC CTGGTAGGGG CATCCCTGT CACCAGGTG CTAAACGACA ACTTTAAACA  5763 AGGCGTGTC CTGGTAGGGG CATCCCTGT CACCAGGTG CTACCACTGT TCAAAATTCC  6764 AGGCGTGTC CTGGTAGGGG CATCCCTGT CACCAGGTG CTGCACACG GGCCACTGA TACGATAGACA GTGGTCCACG GACACTAGC CCCACTGGAC  5762 TGCACACTAC CCTGGTAGCT ATGCGTTCCT CTTGTTTGTA GGGAACGGT GCCCACAGGT  6763 AGGCATCAC TCCCTGGTTCT CTTGTTTGTA GGGAACAGT GCCCACAGGT  6764 AGGCATCAC TACCACAGGA GAACAAACAT CCCTTGCCAC GCCCACAGGAACAA  6764 AATGATATAAGG ATGGGACATA ATGCATACGA GAACAAACAT CCCTTGCCAC GCCACCAGAACACAACAAACAAACAACAAACAAA						
5521 GCTAGATATA TAGCACAGC CTGCCTTCTG CTGCACAAGG CCTTAGAGAC CCGCCCTTTC CGATCTATAT AATCGTCTG GAGGAGAGAC 5581 AATCAGCTTA GCTTGTGCTC TGTTTCTGCT CTGCTTTGG CGAATCTCTG GGCCGAAAG 5581 AATCAGCTTA GCTAGCGAC ACAAAGACGA GAGATCCAG ATTGATTCC TACCCGATAT CGAACACGAG ACAAAGACGA GAGAATCCAG ATTGATTACC ACAGTCAAAA 5561 AATGAACAA AAGTATGCAT CTGCCTTGG CTTGACCTTT TTGCTTTTCA ATGCTAGTATA 5761 AGTGCACAAA AAGTATGCAT CTGCCTTGG CTTGACCTTT TTGGTTTCA ATGCGACTGAA 5761 AGGGGAAAA AAGTATGCAT CTGCCTTGG CTTGACCTT TTGGTTTCA ATGCGACTGAA 5761 AGGGGGAAAA AAGTATGCAT CTGCCTTGC CTTAGCACAAAAACT TACGACTGAAA 5761 AGGGGGAAAA AAGTATCCAC ACACCTTCC CTACCAAAAACT TACGACTGAAA 5761 AGGGGGAAAA AAGTATCCAC CTGCACAGAAACT CATCCCTGT TAAAAATACC AGGGGAAAAAAAAACAAAAAAACAAAAAAAAAA	5461	AGCATTTTTA TTTTATACTC	ATCCAGTGAA	CTCTGCTCTT	CCAAGTGTGT	TCATGTATGT
CGATCTATAT AATCGTGTCG GAGGGAAGAC GAGGTGTTGC GGAATCTCTG GCCGGAAAG  AATGAGCTA GCTTGTCTC TGTTTCTGCT TCTCTAGGTC TAAACTATGG TGTCAGTTTT  TTACTCGAAT GAACAGGAG ACAAAGACGA GAGAATCCAG ATTTGATACC ACAGCCAAAA  5641 AATAGAACAA AAGTATGCAT CTTGCCTTGG CTTGAGCCTT TTCGTTTTCA ATGCTGACTT  TTATCTTGTT TCTCATAGGTA GAACAGGAACC GACCTGGAA AAGCAAAAAGT TACGACTCAAA  5761 CTCCCCTTTC TCTCCTGTGG TCACCTTACC TTTCCAGAGT GTAAAGGGACA ACTTTTAAGG  GAGGGGAAAG AGAGGACACC AGTGCAATGG AAAGGATCAC CATTCCCTGT TGAAAATTCC  TCCGCACAGG GACCATCCC GTAGGGACAC AGTGCCACG CTGTCACCAC CCCCCTTGAG  TCCGCACAGG GACCATCCC GTAGGGACAA GTGGTCACCAC GACGTTGCT GGAAAATTCC  5821 TGACATCTAC CCTGGTGACT ATGCGTTCCT CTTGTTTGTA GGGAACAGTGG GACCATCCCC GTAGGGACAA GTGGTCACCAC GACGTTGCT GGGGGAACACACCTCC CTAGGTGACA TACCCAAGGA GAACAAACAC CCCTTGCACC CAGGGTCACC  5821 TGACATCTAC CCTGGTTGCT TCTGGTTCCC GGCTGCACG GACGTTGCTC CCCCGTGTGACACACACACACACACACACACACACACACA	<b>.</b>					
ANTAGACTTA GCTTGTCTC TGTTTCTGCT CTCTTAGGTC TAAACTATGG TGTCAGTTTT TTACTCGAAT CGAACAGGA CAAAGAGGA GGAATCCAG ATTTGATACC ACAGTCAAAA 5641 ANTAGAACAA AAGTATGCAT CTTGCCTTGG CTTGAGCCTT TTCGTTTTCA TACGTCAGA 5761 CTCCCCTTTC TCTCTGTGC TCACCTTACC TTTCCAGGGT AAAGGGAAAA TACGACAGAA 5761 CTCCCCTTTC TCTCTGTGC TCACCTTACC TTTCCAGGGT GTAAGGGAAAA TACGACTGAA 5761 AGGGGAAAG GAAGGACACA ACTGCAGAAG AAAGCAAAAAT TACGACTGAA 5761 AGGCGTGTCC CTGGTAGGGG CATCCCTGTT CACCAGGTGC CTGTCATCAC CCCACTTGAC 5761 AGGCGTACC CTGGTAGGG CATCCCTGTT CACCAGGTGC CTGTCATCAC CCCACTGAC 5762 AGGCGTACC CTGGTAGGG CATCCCTGTT CACCAGGTGC CTGCAGACAGTG GGCGAAAATACC ACTGTAGAGT GGACCACTCAC TACGCCAGAAA GTGGTCCACG GACCAGTAGTG GGGTGAAATTC 5761 AGGCGTACAC ATCCCTGTTCC TTGTTTTGTA GGGAACAGGTG CGTCCAGGTG ACTGTAGAGT GGACCACTCAA TACCCAAGGA GAACAAAACAT CCCTTGCCAC CGAGGTCCAC 6762 ACTGTAGAGT GGACCACTCAA TACCCAAGGA GAACAAAACAT CCCTTGCCAC CGAGGTCCAC 5771 TGTATATTCC TACCCTACAAT TTGCTTTGTG TGGTGCGAAAACTTT CACAGAAACAT ACACCAAAGGC CCCACCGAAAA CAAAACATT CACAGAACACA ACACCAAGGA CCAACAGAAACAT ACACCAAAACAC ACACCAAGGAAA CAAAAACATT CACAGAACACA ACACCAAGGA CACACAGAAACAT CACACCACAC	5521					
TRACTOGRAT CGRACAGGG RACARAGAGG AGARATCCAS ATTTGATACC ACAGTCCARA  5661 ARTAGARCAR ARGTRGCAT CTTGCCTTGG CTTGAGCCTT TTCGTTTTCR ATGCTGACTT TATCTTGTT TTCATAGGTA GARCGGARCC GARCTCGGAR ARGCGARAG AGAGGACAGG AGTGGARTGG ARAGGTCTAC CATTCCCTGT TACGACTTAC  5761 AGGCGGARAG AGAGGACAGG AGTGGARTGG ARAGGTCTCA CATTCCCTGT TGAARATTCC CCCCACAGG GACCATCCCC GTAGGGG CATCCCTGTT CACCAGGTGC CTGCTCATCAC CCCACTTGAC TCCCCCACAGG GACCATCCCC GTAGGGACAA GTGGTCCACG GACAGTAGTG GGGTAACTG TCCCCACAGG GACCATCCCC GTAGGGACAA GTGGTCCAC GACAGTAGTG GGGTAACTG  5821 TGACATCTAC CCTGGTGACT ATGGGTTCCT CTTGTTTGTA GGGAACGGTG GCTCCAGGGT ACTGTAGATG GGACCACCGA TACCCAAGGA GAACARACAT CCCTTGCCAC CAGAGGTCCAC  5881 GAGGCATCAR TCTGTTGGGT TCTGGTTCCC GGCTGCCCC TGGTTTCAAA CACATATAAGG ATTGGACTCAC AGACCACAGGA GAACARACAT CCCTTGCCAC CAGAGGTCCAC CTCCGTAGTT AGACAACCA AGACCACAGGA GAACARACAT CCACACAGTA CAAAACTTT CACAGAACAC ACATATAAGG ATGGGACCTA TACCCTGCAT TGCTTTGTT GGTTGTAAA GTCTCTTCC CTCCGTAGTT AGACAACCA AGACCACAGGA CCACAGCACTA CCACACACTA ACATATAAGG ATGGGACCTA AACGAACAC ACCACGACTA CCACACACCTA ATCCTAGAAC CTACTGAGAG GTAGTCAGTC TCTGGTGTGC TGTGCAGTAA CCACACACCTA ATCCTAGAACA CTACTGAGAG GTAGTCAGTC TCTGAGGGGG ACAACGTTT ACAGTCCAC TTGCATGGAC CTACTGAGAG GTAGTCAGTG TCTCGAGGGG ACAACGTTT ACAGTCCAC TTGCATGGAC CTACTGAGAG GTAGTCAGTG TCTCGAGGGG ACAACGTTC ACGGCCACCAC TTGCATGGAC CTACCGTAAAA TCTGAGGTGA TTTCTGTTG CGGCTGCACCAC ACGGCCCCC TGCACCGAC CTGGCGATTT ACACCTCAG GGTCGTCCC CGACAGTCGC TGCCCAAAGCT CTACCTTAAA AGGACTGTC CAAAACACAC ACCACGACCCC TGCACCCTAC TAACCATACCAA ACTTCCACTTAC TCCACACGG GTCCTCC GGTTCACCCTAA ACCTACCCTAA ACATGCCAAA CTACGTCAC CAAAACACAC ACCACGACCCCA CTACACCTTAC ACGGACACAC CCCACCACCAC CTACACCTTAC ACGGACACAC CCCACACACCCCAC CTACACCTTAC ACCTCACACG GGTCCACCAC CAAACACAC ACCACCACACAC ACATGCCAAAA CTACACCAC GAGTCACCCAC CAAACACAC ACCACCACACAC ACATGCCAAAA CTACACCAC GGCCCAC CTACACTTCT CACCCTAAAATTC CCACCTCTC TAACCACACAC ACTCACCACACACA						
5611 AATAGAACAA AAGTATGCAT CTTGCCTTGG CTTGAGCCTT TTCGTTTTCA ATGCTGACTT TTACTTGTT TTCATACGTA GAACGGAACC GAACTCGGAA AAGCAAAAGT TACGACTGAA GAGGGGAAAG AGAGGACCG CTACCTTACC CTTCCCAGAGT GTAAGGGACA ACTTTTAACG GAGGGGAAAG AGAGGACACG AGTGGAATGG AAAGGTTCA CATTCCCTGT TAGAAATTCC TCGCCACAGG GACCATCCCC GTAGGGACAG AGTGGTCCCC GTGCATCACC TCGCCACAGG GACCATCCCC GTAGGGACAA GTGGTCCACG GACAGTAGTG GGGTGAACTG ACTGTAGATG GGACCACTCACC GTAGGGACAA GTGGTCCACG GACAGTAGTG GGGTGAACTG S821 RGCATCTAC CCTGGTGACT ATGGGTCCAC GGACAGTAGTG GGGTGAACTG ACTGTAGATG GGACCACTGA TACCCAAGGA GAACAAACAT CCCTTGCCAC CGAGGTCACC ACTGTAGATG GGACCACCAC TACCCAAGGA GAACAAACAT CCCTTGCCAC CGAGGTCCAC CCGTAGAT ACACAACCCA AGACCAAAGGA GAACAAACAT CCCTTGCCAC CGAGGTCCAC CCGTAGATA ACGACAACCCA AGACCAAAGGG CCGACGGGAAA CCAAAACTT CAGAGAAGAG CAATATAAGG ATGGGACGTA AACGAAACAA ACCACGGATAA CCAAAACTT CAGAGAAGAGA	5581	AATGAGCTTA GCTTGTGCTC	TGTTTCTGCT	CTCTTAGGTC	TAAACTATGG	TGTCAGTTTT
THATCHITT TTCATAGGIA GAACGGAACC GAACTCGGAA AAGCAAAGT TACGACTGAA  CTCCCCTTTC TCTCCTGTGC TCACCTTACC TTTCCAGAGGT CTAAGGGACA ACTITITAAGG GAGGGAAAG GAGGGACGG AGTGGAATGG AAAGGTCTCA CATTCCCTGT TGAAAATTCC  TCCCCACAGG GACCACCCC GTAGGGACAG AGTGGCACG CTGTCATCAC CCCACTTGAC TCCCCACAGG GACCACCCC GTAGGGACAG GTGGTCCACG GACAGTAGTG GGGCACACTG ACTGTAGATG GGACCACTGA TAGGGTTCC CTTGTTTGTA GGGAACGGTG GCCCACTGAC ACTGTAGATG GGACCACTGA TAGGGTTCCC GTTGTTTGTA GGGAACGGTG GCCCACGGTCAC  5881 GAGCATCAA TCTGTTGGGT TCTGGTTCC GGTGCCATT GGTTTTGAAA GTCTCTCTC CTCCGTAGTA AGACACCAA AGACCAAAGGG CGACAGGACAAACATT CAGAGAAAGAG CTCCGTAGTA AGACACCAA AGACCAAAGGG CGACAGGAAACAT CCAAAACTTT CAGAGAAAGAG  5941 TGTATATCC CACCAGCAT TAGCTTGTG TGGTGCTGAT CCACACACGT ACAAAACTT CAGAGAAGAG ACATATAAGG ATGGGACCA AGACCACAGGG CGACAGGAAACAC ACCACACGTA CCACACACGT ACCATAGAAC  6001 GATGACTCTC CATCAGTCAC AGACCACAGG CTGTCAGCA CCACACACGT ACCACAGAG TAGCACAGG CTACTGGAAAACAC CACCACACATT CACACTGCACA CTACTGGAAAC  601 CACCGTAAAA TCTGAGTCAG TCACACACAG GCTGTCAGCC ACGGCTTCAA CTACTAGAAC  602 CACCGTAAAA TCTGAGTCAG TCACACACAG GCTGTCAGCC ACGGCTTCAA CTACTAGAAC  603 CACCGTAAAA TCTGAGTCAG TCACACACAG GCTGTCAGCC ACGGCTTCAAACTTC ACAGTCCACACT ATAAAGATAAA AGTGTGCACT CAAAACACA GCCGACACGG TCCCAAAAGGT GAACGTACCG  6121 TATTCTATTT TCACACGTGA GTTCTGTTG CTGGCTGGCT GACTGACATA TAGATACGAT  AAAAGATAAA AGTGTGCACT CAAAAACACA GCCGACCGAA CTGACCGCATA TACTATGCTA  6181 AGTTGAAAT AGGAGTGTC CAAAACACA GCCGACCGAA CTGACCGTAA TATAACAGAT  6241 TGCTCTTTAG TCCTCACACG GGTCGTCCG GTACAGCAC ACTGCCCTTA TACTATGCTAA  6241 TGCTCTTTAG TCCTCACACG GGTCGTCCG GCACTCATCTT CACTGCCGTAA TAGATACAGA  6301 GTGGAAGGAA CAGCACACAG CCACTCATTCT CACTGCCTTATACACACACACACACACACACACACACACA						
THATCHITT TTCATAGGIA GAACGGAACC GAACTCGGAA AAGCAAAGT TACGACTGAA  CTCCCCTTTC TCTCCTGTGC TCACCTTACC TTTCCAGAGGT CTAAGGGACA ACTITITAAGG GAGGGAAAG GAGGGACGG AGTGGAATGG AAAGGTCTCA CATTCCCTGT TGAAAATTCC  TCCCCACAGG GACCACCCC GTAGGGACAG AGTGGCACG CTGTCATCAC CCCACTTGAC TCCCCACAGG GACCACCCC GTAGGGACAG GTGGTCCACG GACAGTAGTG GGGCACACTG ACTGTAGATG GGACCACTGA TAGGGTTCC CTTGTTTGTA GGGAACGGTG GCCCACTGAC ACTGTAGATG GGACCACTGA TAGGGTTCCC GTTGTTTGTA GGGAACGGTG GCCCACGGTCAC  5881 GAGCATCAA TCTGTTGGGT TCTGGTTCC GGTGCCATT GGTTTTGAAA GTCTCTCTC CTCCGTAGTA AGACACCAA AGACCAAAGGG CGACAGGACAAACATT CAGAGAAAGAG CTCCGTAGTA AGACACCAA AGACCAAAGGG CGACAGGAAACAT CCAAAACTTT CAGAGAAAGAG  5941 TGTATATCC CACCAGCAT TAGCTTGTG TGGTGCTGAT CCACACACGT ACAAAACTT CAGAGAAGAG ACATATAAGG ATGGGACCA AGACCACAGGG CGACAGGAAACAC ACCACACGTA CCACACACGT ACCATAGAAC  6001 GATGACTCTC CATCAGTCAC AGACCACAGG CTGTCAGCA CCACACACGT ACCACAGAG TAGCACAGG CTACTGGAAAACAC CACCACACATT CACACTGCACA CTACTGGAAAC  601 CACCGTAAAA TCTGAGTCAG TCACACACAG GCTGTCAGCC ACGGCTTCAA CTACTAGAAC  602 CACCGTAAAA TCTGAGTCAG TCACACACAG GCTGTCAGCC ACGGCTTCAA CTACTAGAAC  603 CACCGTAAAA TCTGAGTCAG TCACACACAG GCTGTCAGCC ACGGCTTCAAACTTC ACAGTCCACACT ATAAAGATAAA AGTGTGCACT CAAAACACA GCCGACACGG TCCCAAAAGGT GAACGTACCG  6121 TATTCTATTT TCACACGTGA GTTCTGTTG CTGGCTGGCT GACTGACATA TAGATACGAT  AAAAGATAAA AGTGTGCACT CAAAAACACA GCCGACCGAA CTGACCGCATA TACTATGCTA  6181 AGTTGAAAT AGGAGTGTC CAAAACACA GCCGACCGAA CTGACCGTAA TATAACAGAT  6241 TGCTCTTTAG TCCTCACACG GGTCGTCCG GTACAGCAC ACTGCCCTTA TACTATGCTAA  6241 TGCTCTTTAG TCCTCACACG GGTCGTCCG GCACTCATCTT CACTGCCGTAA TAGATACAGA  6301 GTGGAAGGAA CAGCACACAG CCACTCATTCT CACTGCCTTATACACACACACACACACACACACACACACA	5641	AATAGAACAA AAGTATGCAT	CTTGCCTTGG	CTTGAGCCTT	TTCGTTTTCA	ATGCTGACTT
GAGGGGAAAG AGAGGACAG AGTGGAATGG AAAGGTTCA CATTCCCTGT TGAAAATTCC AGGCGTGCC CTGGTAGGGG CATCCCTGTT CACCAGGTGC CTGTCATCAC CCCACTTGAC TCCGCACAGG GACCATCCCC GTAGGGACAA GTGGTCCACG GACAGTTGTG GGGTGAACTG TCGCACAGG GACCATCCC GTAGGGACAA GTGGTCCACG GACAGTTGTG GGGTGAACTG TCACCTTAGATG GACCACCAC ATACCCAAGGA GAACAAACAT CCCTTGCCAC CGAGGTCAC CTCCGTAGAT TCTGTTGGGT TCTGGTTCCC GGCTGCCTTT GGTTTTGAAAA GTCTCTTGTC CTCCGTAGATT AGACAACCAA AGACCAAAGGA CCCAAGAACTTT CAGAGAAGAG GACAATAAAGG ATGGGACGTA TAGCTTTGTG TGGTGCAGT GCTGTGGCAG TAGGATCTG ACATATAAGG ATAGGACCAA AACGAAACAC ACCACGACTA CCAAAACTTT CAGAGAAGAG GATGACTCTC CATCAGTCAC AGACTCCCCC TGTTGCAAG TGTCAGGCTG ATCCTAGAAC CTACCTGAGAG GTAGTCAGTG TCTGAGGGGG ACAACGTTA CAGACCGGT AACGTACACC CTACCGTAAAA TCTGAGTCAA GACCACCCC TGTTGCAAG TGTCAGGCTG ACTCGACAG CTACCGTAAAA TCTGAGTCAG TCACACACAG GCTGCCCAAG TACCTAGACG CTACCGTAAAA TCTGAGTCAG TCACACACAG GCTGCCCGC ACGGCTCCA CTGCACAGTCA CTACCTAAAA ACTCGAGTCAG TCACACACAG GCTGCCGC ACGGCTCCA CTGCACAGTCA ATATCATATT TCACCACGTCA GTTCTGTTC CGGCAGGCTG TGCCGAAGGT GAACCTACCG GTGGCATTTT AGACTCACTC AGTGTGTTC CGACAGCCGC TGACCGCAT ATCCTAACGC CTACCTTAAA ACGGAGTGTGC CAAGCACACA GACCACCGC CTGACCGTAA TACATACGAT ATACATATAA AGTGTGCACT CAAAGACAAA GACCACACAC CTGACCGTAA TACATACGAT ACACTTTAGC TCCTCACACG GGTCGTCCG GGTAGTAAGA GTCACCGAAA TTCTTTCG CCACCTTCACAC GGTCGTCCG GGTAGTAAAATT CACTACCAAA GTGCTCACCTAA CAACTCCCAAA CTAGCTACTT GCATAAAATT CGTAAAGTAC CAATGCCAAA CTGCCTACACTAC CACCTTCCGT CCGACACTGG TCAGACGGA CAGCAACAG GTGCACCAAA GTGCCTCAGTA ACATGCCAAA CTAGCTACTT GCATAAAATT CGTAAAGTAC CAATGCCAAA CTGCTTATTATAA ACATGCCAAA CTAGCTACTT GCATAAAATT CGTAAAATTAC CAATGCACAAA CTGCTTATATA ACATGCCAAA CTAGCGACAC GACCGGACAG GAGAAAATA TAATTCATACAA CACCTTGCTT GAACACACA CACCTTGGAACCT TAATTCTTGAA ACCTTGCAT GTACCTCCGG ACTGGAACCA CACCGGGTGC CACAGG GTGAAACCT ACCTTGATA ACCTTGCAT GTACCTCCGG ACTGGAACAC CACCGGGGTG AAACAACAC CACACTGATATA CACACTACAAAAAACAAAAAAAAAAAA		TTATCTTGTT TTCATACGTA	GAACGGAACC	GAACTCGGAA	AAGCAAAAGT	TACGACTGAA
GAGGGGAAAG AGAGGACAG AGTGGAATGG AAAGGTTCA CATTCCCTGT TGAAAATTCC AGGCGTGCC CTGGTAGGGG CATCCCTGTT CACCAGGTGC CTGTCATCAC CCCACTTGAC TCCGCACAGG GACCATCCCC GTAGGGACAA GTGGTCCACG GACAGTTGTG GGGTGAACTG TCGCACAGG GACCATCCC GTAGGGACAA GTGGTCCACG GACAGTTGTG GGGTGAACTG TCACCTTAGATG GACCACCAC ATACCCAAGGA GAACAAACAT CCCTTGCCAC CGAGGTCAC CTCCGTAGAT TCTGTTGGGT TCTGGTTCCC GGCTGCCTTT GGTTTTGAAAA GTCTCTTGTC CTCCGTAGATT AGACAACCAA AGACCAAAGGA CCCAAGAACTTT CAGAGAAGAG GACAATAAAGG ATGGGACGTA TAGCTTTGTG TGGTGCAGT GCTGTGGCAG TAGGATCTG ACATATAAGG ATAGGACCAA AACGAAACAC ACCACGACTA CCAAAACTTT CAGAGAAGAG GATGACTCTC CATCAGTCAC AGACTCCCCC TGTTGCAAG TGTCAGGCTG ATCCTAGAAC CTACCTGAGAG GTAGTCAGTG TCTGAGGGGG ACAACGTTA CAGACCGGT AACGTACACC CTACCGTAAAA TCTGAGTCAA GACCACCCC TGTTGCAAG TGTCAGGCTG ACTCGACAG CTACCGTAAAA TCTGAGTCAG TCACACACAG GCTGCCCAAG TACCTAGACG CTACCGTAAAA TCTGAGTCAG TCACACACAG GCTGCCCGC ACGGCTCCA CTGCACAGTCA CTACCTAAAA ACTCGAGTCAG TCACACACAG GCTGCCGC ACGGCTCCA CTGCACAGTCA ATATCATATT TCACCACGTCA GTTCTGTTC CGGCAGGCTG TGCCGAAGGT GAACCTACCG GTGGCATTTT AGACTCACTC AGTGTGTTC CGACAGCCGC TGACCGCAT ATCCTAACGC CTACCTTAAA ACGGAGTGTGC CAAGCACACA GACCACCGC CTGACCGTAA TACATACGAT ATACATATAA AGTGTGCACT CAAAGACAAA GACCACACAC CTGACCGTAA TACATACGAT ACACTTTAGC TCCTCACACG GGTCGTCCG GGTAGTAAGA GTCACCGAAA TTCTTTCG CCACCTTCACAC GGTCGTCCG GGTAGTAAAATT CACTACCAAA GTGCTCACCTAA CAACTCCCAAA CTAGCTACTT GCATAAAATT CGTAAAGTAC CAATGCCAAA CTGCCTACACTAC CACCTTCCGT CCGACACTGG TCAGACGGA CAGCAACAG GTGCACCAAA GTGCCTCAGTA ACATGCCAAA CTAGCTACTT GCATAAAATT CGTAAAGTAC CAATGCCAAA CTGCTTATTATAA ACATGCCAAA CTAGCTACTT GCATAAAATT CGTAAAATTAC CAATGCACAAA CTGCTTATATA ACATGCCAAA CTAGCGACAC GACCGGACAG GAGAAAATA TAATTCATACAA CACCTTGCTT GAACACACA CACCTTGGAACCT TAATTCTTGAA ACCTTGCAT GTACCTCCGG ACTGGAACCA CACCGGGTGC CACAGG GTGAAACCT ACCTTGATA ACCTTGCAT GTACCTCCGG ACTGGAACAC CACCGGGGTG AAACAACAC CACACTGATATA CACACTACAAAAAACAAAAAAAAAAAA	5701	CTCCCCTTTC TCTCCTGTGC	TCACCTTACC	TTTCCAGAGT	GTAAGGGACA	ACTTTTAAGG
5761 AGGCCTGTCC CTGGTAGGGG CATCCCTGTT CACCAGGTGC CTGCATCAC CCCACTTGAC TCCGCACAGG GACCATCCCC GTAGGGACAG GTGGTCCACG GACCATTACC CTGCATCAC GACCATCACG GACCATCACG GACCATCACG GACCATCACG GACCATCACG CTAGGGACCAG TAGGGTCCACG ACCGTTAGACGTG GACCACTCACACCACCACCACCACCACCACCACCACCACCA		GAGGGGAAAG AGAGGACACG	AGTGGAATGG	AAAGGTCTCA	CATTCCCTGT	TGAAAATTCC
TCCGCACAGG GACCATCCCC GTAGGGACAA GTGGTCCACG GACAGTAGTG GGGTGACTG TGACATCTAC CCTGGTGACT ATGGGTTCCT CTTGTTTGTA GGGAACGATG GCTCCAGGTG ACTGTAGATG GGACCACTGA TACCCAAGGA GAACAAACAT CCCTTGCCAC CGAGGTCCAC  5881 GAGGCATCAA TCTGTTGGGT TCTGGTTCC GGCTGCCTT GGTTTTGAAA GTCTCTTCTC CTCCGTAGTT AGACAACCCA AGACCAAGGG CCGACGGAAA CCAAAACTT CAGAGAAGG  5941 TGTATATTCC TACCCTGCAT TTGCTTTGT TGGTGCTGAT GCTGTGCGA TAGGAACAG ACATATAAGG ATGGGACGTA AACGAAACAC ACCACGACTA CACAGGACTA GACAAACACT ACATATAAGG ATGGGACGTA AACGAAACAC ACCACGACTA CGACACOGTC ATCCTAGAAC  6001 GATGACTCTC CATCAGTCAC AGACTCCCCC TGTTGCAAAG TGTCAGGCCG ACTCGACAGT CTACTGAGAG GTAGTCAGTG TCTGAGGGGG ACACGTTCA ACAGTCCGAC TAGCTAGACC GTGGCATTTT AGACTCAGTC AGTGTGTGC CGACAGTCGA CACAGGACTCAC  6121 TATTCTATTT TCACACGTGA GTTTCTGTTG CTGGCTGG TGCCAAGGT GAACCTACCG 6121 TATTCTATTT TCACACGTGA GTTTCTGTTG CTGGCTGGCT GACCTGAGGT GAACCTACCG 6121 TATTCTATTT TCACACGTGA GTTTCTGTTG CTGGCTGGCT GACCTGAGT GAACCAACG 6121 AATAGAATAA AGTGTGCAC CAAAGACAAC GACCGACCGA CTGACCGTAA TACTATGCTA ATAAGATAAA AGTGTGCAC CCAGACAGAC CCATCATTCT CACTGCTTT GAAACAAAC 6181 ACTGAAATC AGGCGTTGCC CCAGCAGAGC CCATCATTCT CACTGCTTT GAAACAAAC 6241 TGTACGGTTT GATCGATGAA CGTTATTAAA GCATTCATTC CACTGTTTT CAACCAACAG 6241 TGTACGGTTT GATCGATGAA CGTTATTTAAA GCATTCATTC CACTGTTTT CACGGAGAA 6361 TGGAAGGCA GGCTGTGACC AGTCTGCCTG CTCCTTACTA TAATTGTGAG GATTTGTTAC 6301 GTGGAAGGCA GGCTGTGACC AGTCTGCCTG CTCCTTACTA TAATTGTGAG GATTTGTTAC 6361 TGGAACAGTA CATGGAGCAC AGTCTGCCTG GGGCACAGG GTGGAACCTT AGCTCAATT 6361 TGGAACAGTA CATGGAGCAC AGTCTGCTG GGGCACAGG GTGGAACCTT AGCTCAATT 6421 AGTGTGTGT CAAGAGGAA GTCAGGGTAC TAGCTCAAGTA TATTACACCTC CTAAACAATA 6481 TATACATTTG CCCGACACTGG TCACCTGCTG GGGGAACAGG GTGGAACCTT AGCTGAATAT 6481 TATACATTTG CACGACACTG TCACTTGTG GGGCCACAGG GTGGAACCTT AGCTGAATAT 6481 TATACATTTG CCCGTATTAT CTCTAATGTG AACAAAACC CACACTGGGG 651 TGCACACACA AGTTCTCCTT CACTGCATTA TAATTATGG GTTTTCTGTT TCACACACACA ACCACTTTATA 661 TATACATTTG CCCGTATTAT CTCTAATTTG GTTTCTCCATC ATCAGTGTAT TAACTACATA ACCACTGGGG 661 TATCCACTATA AAGGACAAA ACCACTTTTT TCACTGTATA TAACACAAA	5761					
TGACATCTAC CCTGGTGACT ATGGGTTCCT CTTGTTTGTA GGGAACGGTG GCTCCAGGTG ACTGTAGATG GGACCACTGA TACCCAAGGA GAACAAACAT CCCTTGCCAC CGAGGTCCAC S881 GAGGCATCAA TCTGTTGGGT TCTGGTTCCC GCCTGCTCT GGTTTTGAAA GTCTCTCTC CTCCGTAGTT AGACCAACCCA AGACCAAGGG CCGACGGAAA CCAAAACTT CAGAGAAGAG  5941 TGTATATCC TACCCTGCAT TTGCTTTGTG TGGTGCTGAT GCACACGGT TAGGAACAGG ACCATARAAGG ATGGGACGTA AACGAACCA ACCACGACTA CGACACCGTC ATCCTGAGAG ACCATARAAGG ATGGGACGTA AACGAACCAC ACCACGACTA CGACACCGTC ATCCTGAGAG CTACTGAGAG GTAGTCAGTCAC AGACTCCCCC TGTTGCAAAG TGTCAGGCGG TAGGACTCCAC CTACTGAGAG GTAGTCAGTCAG TCTGAGGGGG ACACGTTTC ACAGTCCGAC CTGCGCATTIT AGACTCAGTC AGTTGTTGTTC CTGGCTAGGCC ACGGCTTCCA CTTGCATGGC GTGGCATTIT AGACTCAGTC AGTTGTTGTTC CTGGCTGGCT GACTGCCAAGGT ATAAAGATAAA ACTGTGCACT CAAAGACCAC GCTGCACGCG CACCGCATA ATCTATGCTA ATAAAGATAAAA ACTGTGCACT CAAAGACAAC GACCGACCGA CTGCCGTAAA TATAAGATAAA ACTGTGCACT CAAAGACAAC GACCGACCGA CTGCCGTAAA TTCTATGCTA ATAACGATAAA ACTGTGCACT CAAAGACAAC GACCGACCGA CTGCCGTAA TACTATGCTA ACATCCAAAA CTTGCATGAA CGTATTTAAA GCATTACAT CACTCTCTT GAAACAAACC CAACTTCCGT GACCGACCG CCATCAATTC CACTCTCTT GAACCAACAC CACTTCCCT CAACACCG CCATCAATTC CACTCTCTT CACGCACGTA ACATCGCAAAA CTAGCTTACT GCATAAATT CGTAAAGACA CTTGCTTTC CACACACTA CATCGAGGCA CGCTCTGCC GCCTATACAA CTTGCTTTC CACACCTTCCT CAGCACACTG TCAGCACGAC CACCTTTACTA TAATTGTGAG GATTTCTTAC CACCTTCCT CAGACACTG TCAGACGAC CCCCCTTGCC CACCTTGCAACATAT ACCTTGTCTT CACACACGG TCACCTTGTG GGGCCACAGG GTGGAACCTT ACCTCAATTT ACCACACACA ACTTCCCCGC ACTCGAGTAC TAGCTCACTG CACCTTGATA ACCTTGTCTT CAAAGAGGAA CTCAGGGTAC TAGCACACAG GTGGAACCTT AGCTCAATAT ACCACACACA ACTTCCCCTC ACTGGGTAC TAGCTCACTG CTCAATATCC CACCTTGATA ACCTTGTCT CAAAGAGCAA CTCAGGGTAC TAGCTCACTG CTCAATATCC CACCTTGATA ACCTTGTCT CAAAGAGCAA CTCAGGGTAC TAGCTCACTG CTCAATACTC CACCACACAC ATATCTAAAT ACCTCCTG TATCTCTTT CAGGCTAC TAGCTCACTG CTCAATACC CACCTTGGAA ACCATTGTACTCCTT CAGGGTAC TAGCTCACTG CTCAATACC CACCTTGGAACACACACACACACACACACACACACACACA	•••	TCCGCACAGG GACCATCCCC	GTAGGGACAA	GTGGTCCACG	GACAGTAGTG	GGGTGAACTG
GAGGATCAA TCTGTTGGGT TCTGGTTCC GGCTGCTTT GGTTTTGAA GTCTCTTCTC CTCCGTAGTT AGACAACCA AGACCAAGGG CCGACGGAAA CCAAAACTTT CAGAGAAGAGA	5821					
GAGGCATCAA TCTGTTGGGT TCTGGTTCCC GGCTGCCTTT GGTTTGAAA GTCTCTTCTC CTCGTAGTT AGACAACCCA AGACCAAGGG CCGACGGAAA CCAAAACTTT CAGAGAAGGG CCGACGGAAA CCAAAACTTT CAGAGAAGAGG ACATATAAGG ATGCGACGTT TTGCTTTGTG TGGTGCTGAT GCTGTGGCC ATGCATCTTG ACATATAAGG ATGCGACGTA AACGAAACAC ACCACGACTA CCACACGGCT ATCCATCAGACG CTACTGAAGG GTAGTCAGG AACGCACACAC ACCACGACTA CCACACGGCT ATCCATCAGACG CTACTGAGAG GTAGTCAGT TCTGAGGGGG ACAACGTTT ACAGTCGCA CTACACACGG GTTGCAAAA TCTGAGTCAG TCTCACACACAG GCTGTCAGC ACGCCTTCCA CTTGCATGCC GTGCGCATTTT AGACTCAGTC ACTGTGTGC CGACAGCGG CTCCGAAGGT GAACGTACCG ATAAACATATA AGGTGCACT CAAAGACAAC GCCCACCGG CTCACCGTAA TACTATGCTA ATAAGATAAA AGTGTGCACT CAAAGACAAC GACCGACCGG CTGACCGTAA TACTATGCTA ATAAGATAAA AGTGTGCACT CAAAGACAAC GACCGACCGG CTGACCGTAA TACTATGCTA ACATGCCAAA CTGCACACAC GGTCGTCTCG GGTAGATAAAATT CACATGCCAAAA CTTGTTTCGTTC GAAAGACAAC CTACTTCACTT GAAACAAACC ACCATCATCT CACTGTCTT GAAACAAACC ACCATCAACAC CACCTTCCAACAC GGTCGTCTCG GGTAGATAAAATT CACATGCCAAAA CTACGTACATA CATGCACAAA CTACGTACAT ACATACACACAA GTGCCCAAAAATT CGTAAAATTT CGTAAAATTA CACATGCCAAAA CTACGTACAT ACATGCACAAA CTACGTACAT ACACTTCCGG ACACCATTG CACCTTCCGGA ACACCTTCT CACGACACAT ACACTTCATA ACACTTCCAT GAACACACG ACACCTTCTGGA ACACCTTCACACACACAG AGATTCCCCC ACCCGCGTGTCC CACCCTTGGAA TCGACTATAA ACCTTGTACAC GACATTATACAC CCCCCTGTCC CACCTTGGAA TCGACTTATAA ACCTTGACACACAG AGATTACAC TCACATATACAC TTTATTATCGG GTGTACCTACAAAATAC CACCTTGGAA TACACTTATAA ACCTTGAACAAAAAACACAAAAACACAAAAACAAAAACAAAAACAAAA		ACTGTAGATG GGACCACTGA	TACCCAAGGA	GAACAAACAT	CCCTTGCCAC	CGAGGTCCAC
TCACGTAGTT AGACAACCCA AGACCAAGGG CGGACGGAAA CCAAAACTT CAGAGAAGAG  5941 TGTATATTCC TACCCTGCAT TTGCTTTGT TGGTGCTGAT GCTGTGGAG TAGGATCTTG ACATATAAGG ATGGGACGTA AACGAAACAC ACCACGACTA CGCACACGTC ATCCTGAGAC  6001 GATGACTCTC CATCAGTCAC AGACTCCCC TGTTGCAAAG TGTCAGGCTG ACTCGACAGT CTACTGAGAG GTAGTCAGT TCTGAGGGGG ACAACGTTTC ACAGTCCGAC TGAGCTGTCA  6061 CACCGTAAAA TCTGAGTCAG TCACACACAG GCTGTCAGCC ACGCCTTCCA CTTGCATGCC GTGCCATTTT ACACGTGA GTTTCTGTTG CGACAGTCGG TGCCGAAGGT GAACGTACCG  6121 TATTCTATTT TCACACGTGA GTTTCTGTTG CTGGCTGGCT GACCTGCATT ATCTATGCTA ATAAGATAAA AGTGTGCACT CAAAGACAAC GACCACCGA CTGACCGTAA TAGATACGAT ATAAGATAAA AGTGTGCACT CAAAGACAAC GACCACCGA CTGACCGTAA TAGATACGAT CAACTTTAG TCCTCACACG GGTCGTCCG GGTAGTAGA GTGACCAGAAC  6241 TGTACGGTTT GATCGATCAG CGTATTTAAA CGATTTCATG CAATGACAAAA CTTGTTTCG ACAGCCCAAA CTAGCACACG GGTCGTCCG GGTAGTAAGA GTGACCAGAA CTTGTTTCTTCA ACATGCCAAAA CTAGCATCAA CGTATTTAAA CGATTCATG CAATGACAAAA CTTGTTTCAG ACATGCCAAAA CTAGCATCAG GACCTGCCT CCCTAAAGTAC GTACTGTTT CACCTCCATA CATGGAGCC TCAGCTTGTG GGGCACACGG GTGCAACCTT ACCTAGTAA ACCTTCCGT CCGACACTGG TCAGACCGGAC GAGGAATGAT ATTAACACTC CTAAACAATG  6301 TGGAACAGTA CATGGAGGCC TGACCTTGTG GGGCCACAGG GTGCAACCTT ACCTGAATATA ACCTTGCAT CATGGAGGCC TCAGCCTTGTG GGGCCACAGG GTGCAACCTT ACCTGAATATA ACCTTGCAT CAAGAGGGAA GTCAGGGTAC TAGCTGAGCT ACCTGGAATATA ACCTTGCAT CAAGAGGGAA GTCAGGGTAC TAGCTCAGTATA ACCTTGCAT CAAGAGGGAA GTCAGGGTAC TAGCTCAGTC CACCATTGGATTA ACCTTGCAT CAAAGAACATA CATGTAGACC CCCCGTGTCC CACCATTGGAT TCCACACACAG AGTTCTCCTT CAGTCCCATG ATCGAGTCAC GAGTTAGAG TCCATTATA ACCTTGCACCAAAAAAAACATT CTCAATATGG GAGTCCCCCAC TTCTTTTTTTAGG GTTTTATCTGTG ATAGCGTACCT AAAAGACCAT ATCAATATACC CACAGGGGTG AAAGAACCAA ACCAGTGGGG  6601 GATCCCCCGG TCTTCTGCTG TATCTAGAAC AGTGACCATATA AATGATTGTAT GGGAAAAAAAAACCAA AACAACATT TCTATTATAGG GTTTATCTACAAAAAACAC TTCAATATACC CACAAGGGTG AAAAAAAAACACAAAAAAAAAA	5881					
TOTATATTCC TACCCTGCAT TTGCTTTGTG TGGTGCTGAT GCTGTGGCAG TAGGATCTTG ACATATAAGG ATGGACGTA AACGAAACAC ACCACGACTA CGACACGGTC ATCCTAGAAC GTGACTCTC CATCAGGTCAC AGACTCCCC TGTTGCAAAG TGTCAGGGTG ACTCGACAGT CTACTGAGAG GTAGTCAGTG TCTGAGGGGG ACAACGTTC ACAGTCCGAC TGAGCTGTCA GCACGATAAA TCTGAGTCAG TCTCACACACAG GCTGTCAGCC ACAGTCCGACTT AGACTCAGCC GTGGCATTTT AGACTCAGTC AGTGTGTCC CGACAGTCGC ACAGCCGACGT GAACGTACCG GTGGCATTTT TCACACGTGA GTTTCTGTTG CTGGCTGGCT GACCGCAATA ATCTATGCTA ATAAGATAAA AGTGTGCACT CAAAGACAAC GACCGACCGA CGACCGATA ATCTATGCTA ATAAGATAAA AGTGTGCACT CAAAGACAAC GACCGACCGA CTGACCGATA ATCTATGCTA ACATGCCAAA CAGGAGTGCC CCAACAGACC CACCACACAGA CTGACCGATA ACATGCCAAA CTACGTTGACACTT GCATAAAATT CGTAAAGAACA GTGCACAGAA CTTTGTTTCG  G241 TGTACGGTTT GATCGATCAA CGTATTTAAA GCATTCATG CAATGACAAAA CTTTGTTTCG CACCTTCCGT CAGCACACG GGTCGTCTCG GTAACAAAGA GTGCTCAGTA ACATGCCAAA CTAGCTGCT CAGACAGAC CTCCTTCATA TAATTGTGAA CACCTTCCGT CCGACACTGG TCAGCACGAC GAGGAATGAT ATTAACACTC CTAAACAATG CACCTTCCGT CCGACACTGG TCAGCCTTGT CTCCTTACCTA TAATTACACTC CTAAACAATG ACCTTGTCAT GTACCTCCGG ACTGGAACAC CCCCGTGTCC CACCTTGGAA TCGACTATA ACCTTGTCAT GTACCTCCGG ACTGGAACAC CCCCGTGTCC CACCTTGGAA TCGACTATA ACCTTGTCAT GTACCTCCGG ACTGGAACAC CCCCGTGTCC CACCTTGGAA TCGACTATA TCACACACAG AGTTCTCCTT CAGGCGCAC ATCAGTCAC GAGTTACAGG TCCATCATAA ATAGTAAAC GGGCAAAATA GAGATTACAC TTTATTTAGG GGTTTGGAG TCCAAGACAC  6541 TATACATTG CCCGTTTTAT CTCTAATATGG GGTTTGTGAG CAAACACCT GTTTATCGTG ATACGATACA AGAGCGAA AGGATTACAC TTTATTTAGG GGTTTGGAA ACAATAGACC  6541 TAGCGTACCT AAAAGACTAT TCTATTATGG GGTTTGTAAA ACAATAGACC  6541 TAGCGTACCT AAAAGACTAT TCTATTATGG GGTTTGCACC TTTTTTTGGT TCACCACCGG TCTTCTGCTG TATCTATAACAC CACAGGGGTG AAAGAACCAA ACCAGTGGGG  6601 GATCCCCCGG TCTTCTGCTG TATCTATAACAC CACAGGGGTG AAAGAACCAA ACCAGTGGGG  661 TTTCCATATG ATCTGTTC TGGAGTATAT CCTACAGTAT TAACTACCAC AAAAAACCAC TTTTTTTGGG TTTTTTTGGG TTTTTTTTTTTTTTTT	5001	CTCCGTAGTT AGACAACCCA	AGACCAAGGG	CCGACGGAAA	CCAAAACTTT	CAGAGAAGAG
ACATATARAGG ATGGGACGTA ANCGARACAC ACCACGACTA CGACACCGTC ATCCTAGAAC GATGACTCTC CATCAGTCAC AGACTCCCCC TGTTGCARAGG TCTCAGGCTG ACTGACCAGT CTACTGAGAG GTAGTCAGT TCTGAGGGGG ACAACGTTTC ACAGTCCGAC TGAGCTGTCA GCACGTARAR TCTGAGTCAG TCACACACAG GCTGTCAGCC ACGGCTTCCA CTTGCATGCC GTGGCATTTT AGACTCAGTC AGTGTGTGTC CGACAGTCGG TGCCGAAGGT GAACGTACCG ATAACATAAA AGTGTGCACC CAAAGACAAC GACCGACCGA CTGACCGTAA TAGATACGAT ATAACATAAA AGTGTGCACC CAAAGACAAC GACCGACCGA CTGACCGTAA TAGATACGAT ACAACTTAG TCCTCACACG GGTCGTCTCG GGTAGTAAGA CTTGTTTCG TCAACCTTAG TCCTCACACG GGTCGTCTCG GGTAGTAAGA CTTGTTCTG ACATGCCAAA CTAGCTACTT GCATAAATTT CGTAAACTAC GTGACCAAA CTTGTTCTG GGGGAAGGCA GGCTGGACCA CTCACTGTCTT CACGAGTCAT ACATGCCAAA CTAGCTACTT GCATAAATTT CGTAAACTAC GTTACTGTTT CACGAGTCAT ACATGCCAAA CTAGCTACTT GCATAAATTT CGTAAACTAC GTTACTGTTT CACGAGTCAT ACATGCCACAC CGGCACCTGG TCAGACCGAC CGCGAACACAT ATTAGTGAG GATTGTTAC CACCTTCCGT CCGACACTGG TCAGACCGAC CGCGAACACAT AATTAGCACC CTAAACAATG ACCTTGCAT CATGGAGGCC TCAGACCGAC CCCCGTGTCC CACCTTGGAA ACCTTGTCAT GTACCTCCGG ACTGGAACAC CCCCGTGTCC CACCTTGGAA ACCTTGTCAT GTACCTCCGG ACTGGAACAC CCCCGTGTCC CACCTTGGAA ACCTTGTCAT GTACCTCCGG ACTGGAACAC CCCCGTGTCC CACCTTGGAA ACCTTGTCAT CAAGAGGAA GTCAGGGTAC TAGCTCAGTG CTCAATCACC ACCTTGCAT GTACCTCCGT ACTGCACCAC ACCCTTGGAA TATTACACACC ACCTTGAACATA GAGTTCCCTT CAGTCCCATG ATCGACTCAC GAGTTACGAG ACCTTGTCAT CACACACAG AGTTCTCCTT CAGTCCCATG ATCGACTCAC CACCTTGGAA TATTACCACACACACACACT TTTATCTGCTG ATATGTAAAC GGGCAAAATA GAGATTATACC CACAGGGGTG AAAGAACACT TTTATCGTG ATATGTAAAC GGGCAAAATA GAGATTATACC CACAGGGGTG AAAGAACCAA ACCAGTGGGG  GATCCCCCGG TCTTCTGGTG TATCTATTATGG GTGCCCCAC TTTTTTTGGT TTTATCGGG ACTGCGCACT AAAACACTAT TCTTATTATGG GTGCCCCAC TTTTTTTTTAGG GTTTTTTCGGG ACCCCCGG TCTTCTGGTT TCTTATTATGG TACACTATA AATGATATA CCCTTATCAC  6601 GATCCCCCGG TCTTCTGGTG TATCTATATACC CACAGGGGTG AAAGAACCA ACCAGTGGGG AAAGGACACA ACCATGTT CTTATATAC GGGAATAATA CCCTTATCAC  661 TTTCCATATG ATCTGTTTC TGGAGTTATA CACACATGT CATTTACTGT ACAAAAACCC AAAGGTATAC TAGAACAAAAACAG ACCTCATATA CACACATGT CACGCCCCC TATTTTACAG  6721 ACGCGCACC AGAACAACA	5941					
GATGACTCTC CATCAGTCAC AGACTCCCCC TGTTGCAAAG TGTCAGGCTG ACTGGACAGT CTACTGAGAG GTAGTCAGT TCTGAGGGGG ACAACGTTTC ACAGTCCGAC TGAGCTGTCA GTGCTGAAAA TCTGAGTCAG TCACACACAG GCTGTCAGCC ACGGCTTCCA CTTGCATGGC GTGGCATTTT AGACTCAGT AGACTCAGT CGACAGTAGG TGCCGAAAGT GAACGTACCG GTGGCATTTT AGACTCAGT AGTGTGTGTC CGACAGTCGG TGCCGAAAGT GAACGTACCG GTGGCATTTT TAGACTCAGT ATTCTGTTG CTGGCTGGCT GACCGGAAGT GAACGTACCG ATAACATAAA AGTGTGCACT CAAAGACAAC GACCGACCGA CTGACCGTAA TAGATACGATAAA AGTGTGCACT CAAAGACAAC GACCGACCGA CTGACCGTAA TAGATACGATAAA AGTGTGCACT CAAAGACAAC GACCGACCGA CTGACCGTAA TAGATACGAT TAAACATAAA AGTGTGCACC CCAGCAGAGC CTGACCGTAA TAGATACGAT AACATCTAGG TCCACACACG GGTCGTCTCG GGTAGTAAGA CTAGCTCAGTA TAGATACAAAA CTTTGTTTCC GACTACTATA ACATCCCAAA CTAGCTACTT GCATAAATTT CGTAAACTAC GTGACCAAAA CTTTGTTTCC CACCTCCGT CACCTTCCGT CACCACTCCGT CACACACAC GGTCGCCGT CTCCTTACTA TAATTGTGAG GATTTGTTAC CACCTCTCCGT CAGACACAC GACCTTCCGT CAGACCACA GACCACACACACACACACACACACACACA	3342.	ACATATAAGG ATGGGACGTA	AACGAAACAC	ACCACGACTA	CGACACOGTC	ATCCTAGAAC
CTACTGRAGA GTAGTCAGTG TCTGAGGGGG ACAACGTTTC ACAGTCCAC TGAGCTGTCA  6061 CACCGTAAAA TCTGAGTCAG TCACACACAG GCTGTCAGCC ACGGCTTCCA CTTGCATGGC GTGGCATTTT AGACTCAGTC AGTGTGTC CGACAGTCG TGCCGAAGGT GAACGTACCG  6121 TATTCTATTT TCACACGTGA GTTTCTGTTG CTGGCTGCGC TGCCGAAGGT ATCTATGCTA ATAAGATAAA AGTGTGCACT CAAAGACAAC GACCGACCGA CTGACCGTAA TAGATACGAT  6181 AGTTGAAATC AGGAGTGTGC CCAGCAGAGC CCATCATTCT CACTGTCTTT GAAACAAAGC  6241 TGTACGGTTT GATCGATGAA CGTATTTAAA GCATTTCATG CAATGACAAA GTGCTCACTA ACATGCCAAA CTAGCTACTT GCATAAATTT CGTAAAGTAC GTTACTGTTT CACGAGTCAT  6301 GTGGAAGGCA GGCTGTGACC AGTCTGCCTG CTCCTTACTA TAATTGTGAG GATTTGTTAC CACCTTCCGT CCGACACTGG TCAGACGGAC GAGGAATGAT ATTATACACTC CTAAACAATG  6361 TGGAACAGTA CATGGAGGCC TGACCTGTG GGGCACAGG GTGGAACCTT AGCTGAATAT ACATGCTAGT CTACCTCCGG ACTGGAACAC CCCCGTGTCC CACCTTGGAA TCGACTATA ACCTTGTGTC TCAAGAGGAA GTCAGGGTAC TAGCTCAGTG CTCAATCTCC AGGTACTATA ACCTTGTGT GTACCTCCGG ACTGGAACAC CCCCGTGTCC CACCTTGGAA TCGACTATA ACCTTGTGT CTCAAGAGGAA GTCAGGGTAC TAGCTCAGTG CTCAATCTCC AGGTACTATA ACCTGTGTGT CTCAAGAGGAA GTCAGGGTAC TAGCTCAGTG CTCAATCTCC AGGTACTATA ACCTGTGTGT CTCAAGAGGAA GTCAGGGTAC TAGCTCACTG CTCAATCTCC AGGTACTATA ACCTGTGTGT CTCAAGAGGAA GTCAGGGTAC TAGCTCACTG CCACCTTTGAA TCGACCACAG AGTTTCTCTT CAGTCCCATG ATCGAGTCAC CACCTTGGAA TCGACTATA  6421 AGTGTGTGT CTCAAGAGGAA GTCAGGGTAC TAGCTCACTG CTCAATCTCC AGGTACTATA ACTGCACACA AGTTCTCCTT CAGTCCCATG ATCGAGTCAC CACAGAGCT TTTTTTCTGGT ATACCATCAG AGTTCTCCTT CAGTCCCATG AAAAAAAACC CCAAACACTT GTTTATCCTG ATATGTAAAC GGGCAAAAATA GAGATATACAC TTTATTTAGG GGTTTTGGAA CAAATAGCAC  6541 TAGCCTACCT AAAAGACTAT TCTATATAGG GTGTCCCCAC TTTTTTTGGTT TGGTCACCCC ATCGGGGGC AGAAGACGAC ATAGAATCTT TCACTGATAT TTACTTGATA AATGAACCAA ACCAGTGGGG  6601 GATCCCCCGG TCTTCTGCTG TATCTAGAAC AGCAGGGTG AAAGAACCAA ACCAGTGGGG AAAGGAACAAT ATCACTTTTT TGGAGTATAT CTCATATATA CCTCAGATAT TATCTTACTAT AAAAAACCC AAAGGAACAAT ATCACTTTTT TGGAGTATAT TTACTTACTAT AAAAAACCC AAAGGTATAC TAGACAAAA ACCAGTTTTC TCTCTGTTT CACCTGCAC TATTTAAAAA TCACGTCGAC TAGACAAAA ACCAGTTTCT TCTTGTAAA GAAAAAACG CAGAATATTTT AAGGGTACCAAAAACAC C	6001	CATGACTOTO CATCAGTOAC	AGACTCCCCC	TGTTGCAAAG	TGTCAGGCTG	ACTOGACAGT
6061 CACCGTARAR TCTGAGTCAG TCACACACAG GCTGTCAGCC ACGGCTTCA CTTGCATGGC GTGGCATTT AGACTCAGTC AGTGTGTC CGACAGTCG TGCCGAAGGT GAACGTACCG 6121 TATTCTATTT TCACCAGTGA GTTTCTGTG CTGGCTGGC TGCCGAAGGT ATCATAGCTA ATAAGATARA AGTGTGCACT CARAGACAAC GACCGACCGA CTGACCGTAA TAGATACCAT 6181 AGTTGAAATC AGGAGTGTC CCAGCAGAGC CCATCATTCT CACTGTTTT GAAACARAGC TCAACTTTAG TCCTCACACG GGTCGTCTCG GGTAGTAAGA GTGACAGAA CTTTGTTTCG 6241 TGTACGGTTT GATCGATGAA CGTATTTAAA GCATTTCATG CAATGACAAA GTGCTCAGTA ACATGCCAAA CTAGCTACTT GCATAAATTT CGTAAAGTAC GTTACTGTTT CACGAGTCAT ACATGCCAAA CTAGCTACTT GCATAAATTT CGTAAAGTAC GTTACTGTTT CACGAGTCAT CACCTTCCGT CCGACACTGG TCAGACGGAC CACCGTACTACTA TAATTGTGAG GATTTGTTAC CACCTTCCGT CCGACACTGG TCAGACGGAC GAGGAATGAT ATTACACACTC CTAAACAATG CACCTTCCAT GTACCTCCGG ACTGGAACAC CCCCGTGTCC CACCCTTGGAA TCGACTTATA ACCTTGTAT GTACCTCCGG ACTGGAACAC CCCCGTGTCC CACCCTTGGAA TCGACTTATA ACCTTGTAT TCAAGAGGAA GTCAGGGTAC TAGCTCAGTG CTCAAACACT AGCTCATATA CCACACACAG AGTTCTCCTT CAGTCCCATG ATCGAGTCAC GAGTTAGAGG TCCATGATAT TCACACACAG AGTTCTCCTT CAGTCCCATG ATCGAGTCAC GAGTTAGAGG TCCATGATAT TCACACACAG AGTTCTCCTT CAGTCCCATG ATCGAGTCAC GAGTTAGAGG TCCATGATAT TCACACACAG AGTTCTCCTT CAGTCCCATG ATCGAGTCAC CACAGTGGG TAGAGACCA ACCAGTGGGG 6541 TAGCGTACCT AAAAGACTAT TCTATATAGG GTGTCCCCAC TTCTTGGTT TGGTCACCCC ATCGCATGGA TTTTCTGATA AGATATACC CACAGGGGTG AAAGAACCAA ACCAGTGGGG 6601 GATCCCCCGG TCTTCTGGTG TATCTTATAGG GTGTCCCCAC TTCTTGTGTT TGGTCACCCC AAAGGAACCAA ACCAGTGGGG 661 TTTCCATATG ATCTGTTCT TGGAGTATAT CCTACATGTT CACTGGAGAAAAAACCC CTAGGGGGCC AGAAGACCAA ACCAGTTGGTT TAGAGACAAAACCC AAAGGATAAC TAGATCTTG TCACTATATA GCTACATATA AATGATGATA CCCTTATCAC 6661 TTTCCATATG ATCTGTTCT TGGAGTATAT CCTACATGTT CACTGGAGGG AAAGAACCAA ACCAGTGGGG 6721 AGGGGGCC AGAAGAACAAG ACCTCATATA CGATGTATA TTACTACATA CCCTTATCAC 6721 AGGGGGCC AGAAGAACAAG ACCTCATATA CGATGTATA GTACATATA CCCTTATAAAATTTTT CACCGTCGAC TAGACCAAAA ACCAGTGTGT CTCTGTTATACTG TACACTATAA TTATTTGGG 6721 AGGGCACATG ATGATCACT TCTGTTTATAAAAATTTTT CACCGTCGAC TACACCGTATAAA GCACTCTCT TCTGTTTATA GAAAAAACG CAGCGTCTGG	0001	CTACTGAGAG GTAGTCAGTG	TCTGAGGGGG	ACAACGTTTC	ACAGTCCGAC	TGAGCTGTCA
6121 TATTCTATT TCACACGTCA GTTTCTGTG CTGGCTGGCT GACCGAAGGT GAACGTACCG 6121 TATTCTATTT TCACACGTGA GTTTCTGTG CTGGCTGGCT GACCGGAAT ATCTATGCTA ATAAGATAAA AGTGTGCACT CAAAGACAAC GACCGACCGA CTGACCGTAA TAGATACGAT 6181 AGTTGAAATC AGGAGTGTGC CCAGCAGAGC CCACATTCT CACTGTCTTT GAAACAAAGC 6241 TCTACGGTTT GATCGATGAA CGTTCTTCGGGTTAGAAGAGAAA CTTTGTTTCG 6241 TGTACGGATTA GACCGACACA CGACCGACGA CGAAGAAAA GTGCTCAGTA 6301 GTGGAAGGAA CTAGCTACAT GCATAAAATT CGTAAAAGAA GTGCCCAGAA CACCTTCCGT CCGACACTGG TCAGACCGAC GAGGAATGAT ATTAACACTC CTAAACAATG 6361 TGGAACAGTA CATGGAGGCC TGACCTTGTG GGGGCACAGG GTGGAACCTT AGCTGAATAT ACCTTGTCAT GTACCTCCGG ACTGGAACAC CCCCGTGTCC CACCTTGGAA TCGACTTATA 6421 AGTGTGTGT TCAAGAGGAA GTCAGGGTAC TAGCTCAGTG CTCAATCTAC AGGTACTATA TCACACACAG AGTTCTCCTT CAGTCCATG ATCGACTAC GAGTTAGAGG TCCATGATAT TCACACACAG AGTTCTCCTT CAGTCCCATG ATCGACTAC GAGTTAGAGG TCCATGATAT 6481 TATACATTTG CCCGTTTTAT CTCTAATGTG AAATAAAATCC CCAAACACTT GTTTATCGTG 6541 TAGCGTACCT AAAAGACTAT TCTTATATGG GTTTCGTGAA CAAATAGCAC 6541 TAGCGTACCT AAAAGACTAT TCTTATATAGG GTTTTGTGAA CAAATAGCAC 6641 TAGCGTACCA AAAAGACTAT TCTTATATAGG GTTTCTGTAA CAAATAGCAC 6661 TAGCGTACGA TTTTCTCATA AGATTAACAC TTTTTTTTGG GGTTTTGTGAA ACCAGTGGGG 6661 TTTCCATATG AAGAACAAA ACCAGTGGGG 6661 TTTCCATATG AAAAAAAAA GAAATAATAC CACAGGGGTG AAAGAACCAA ACCAGTGGGG 6661 TTTCCATATG ACAACACAA ACCAGTGGGG 6721 AGGGCACCAAAAAAA ACCAGTTCTCTC TCTGTATATA CAAAAAAAAAA	6061	CACCETAAAA TCTGAGTCAG	TCACACACAG	GCTGTÇAGCC	ACGGCTTCCA	CTTGCATGGC
ATATCTATTT TCACACGTGA GTTTCTGTTG CTGGCTGGCT GACTGGCATT ATCTATGCTA ATAAGATAAA AGTGTGCACT CAAAGACAAC GACCGACGA CTGACCGTAA TAGATACGAT  AGTTGAAATC AGGAGTGTGC CCAGCAGAGC CCATCATTCT CACTGTCTTT GAAACAAGC TCAACTTTAG TCCTCACACG GGTCGTCTCG GGTAGTAAGA GTGACAGAAA CTTTGTTTCG  6241 TGTACGGTTT GATCGATGAA GCTATTTAAA GCATTTCATG GAAACAAAA GTGCCAGTA ACATGCCAAA CTAGCTACTT GCATAAATT CGTAATGACAC GTTACTGTT CACGAGTCAT  6301 GTGGAAGGCA GGCTGTGACC AGTCTGCTG CTCCTTACTA TAATTGTGAG GATTTGTTAC CACCTTCCGT CCGACACTGG TCAGACGGAC GAGGAATGAT ATTAACACTC CTAAACAATG  6361 TGGAACGATA CATGGAGGCC TGACCTTGTG GGGCACAGG GTGGAACCTT AGCTGAATAT ACCTTGTCAT GTACCTCCGG ACTGGAACAC CCCCGTGTCC CACCTTGGAA TCGACTTATA TCACACACAG AGTTCTCCTT CAGTGCACT ATCGAGTCAC GAGTTAGAGG TCCAATATT TCACACACAG AGTTCTCCTT CAGTCCCATG ATCGAGTCAC CAAACACACT GTTTATCGTG ATATGTAAAC GGGCAAAATA GAGATTACAC TTTATTATGG GGTTTGTAACACCC ATAGCGAGCA TTTTCTGATA AGATTAATACC CACAGGGGTG AAAGAACCAA ACCAGTGGGG  6601 GATCCCCCGG TCTTCTGCTG TATCTATGTG GTGTCCCCAC TTTCTTGGTT TGGTCACCCC ATGGGGGCC AGAAGACACAC ATAGATCTTG TCACTGATAT TTACTTACATA CTAGGGGGCC AGAAGACGAC ATAGATCTTG TCACTGATAT TTACTTACATA CTAGGGGGCC AGAAGACGAC ATAGATCTTG TCACTGATAT TTACTTACATA CTAGGGGGCC AGAAGACGAC ATAGATCTTG TCACTGATAT TTACTTACATA CACGCATGGA TTTTCTGGTT TATCTGATA GGTGACCACA ACCAGTGGGG  6661 TTTCCATATG ATCTGTTCT TGGAGTATAT CACTGATATA AATGATCACAC ACCAGTGGGG  6721 AGGGCACCAC AGAAGACGAC ATAGATCTTG TCACTGATAT AAAGAACACA ACCAGTGGGG  6721 AGGCCACCAC AGAAGACGAC ATAGATCTTG TCACTGATAT AAAGAACCAC TTTATCTGCAC AAAAGAACACA ACCAGTGGGG  6721 AGGCCACCAC AGAAGACAACA ACCTCATATA CCACTGATATA CATTTACTGT AAAAAAACCC AAAAGGTATAC TAGACAACAG ACCTCATATA CCACTGATATA AATGAACAC TTTTTTCTGGT AAAGGTATAC TAGACAACAG ACCTCATATA CCACTGATATA AATGAACA TTTTTTTAGGG  6721 AGGCCCCAC TACTACGTTT CTGGAGAAACAG ACCTCATATAA CCACTGTTACAC TTTTTTTTTT	0001	GTGGCATTTT AGACTCAGTC	AGTGTGTGTC	CGACAGTCGG	TGCCGAAGGT	GAACGTACCG
ATARGATARA AGTGTGCACT CARAGACARC GACCGACCGA CTGACCGTAR TAGATACGAT  6181 AGTTGARATC AGGAGTGTGC CCAGCAGAGC CCATCATTCT CACCTCTTT GARACARAGC  TCARCTTAG TCCTCACAGG GGTCGTCTG GGTAGTARA GTGACAGARA CTTTGTTTCG  6241 TGTACGGTTT GATCGATGAR CGTATATARA GCATTTCATG CARTGACARA CTTGCTTTGT  ACATGCCARA CTAGCTACTT GCATARATTT CGTARAGTAC GTTACTGTT CACCAGAGCAT  6301 GTGGARGGCA GGCTGTGACC AGTCTGCCTG CCACCTTCCGT CCGACACTGG TCAGACGAC GAGGARGAT ATTARACACTC CTARACATG  CACCTTCCGT CCGACACTGG TCAGACGGAC GAGGARGAT ATTARACACTC CTARACATG  ACCTTGTCAT CACACAGGA ACTCCCCTGTGTG GGGCCACAGG GTGGARCCTT AGCTCAATAT  ACCTTGTCAT CTARAGAGGAR GTCAGGGTAC TAGCTCAGTG CTCATATCACACACAG ACTCCCCTCT CAGTCCCATG ATCGAGTCAC GAGTTAGAGG TCCATGATAT  6421 AGTGTGTCT CTARAGAGGAR GTCAGGGTAC TAGCTCAGTG CTCAATCTCC AGGTACTATA  6421 TATACATTTG CCCGTTTTAT CTCTAATGTG AAATRAATCC CAAACACTT GTTTATCGTG  ATATGTAAAC GGGCAAAATA GAGATTACAC TTTATTTAGG GGTTTGGAA CCAAATAGCAC  6541 TAGCGTACCT AAAAGACTAT TCTATTATGG GTGTCCCCAC TTTCTTGGTT TGGTCACCCC  ATAGGTGGA TTTTCTGATA AGATAATACC CACAGGGGTG AAAGAACCAA ACCAGTGGGG  6601 GATCCCCCGG TCTTCTGCTG TATCTAGAAC AGTGACTATA AATGATGTAT GGGAATAGTG  CTAGGGGGCC AGAAGACGAC ATAGATCTTG TCACTGATAT TTACTACATA CCCCTTATCAC  6661 TTTCCATATG ATCTGTTGTC TGGAGTATAT GCTACATATA TTACTACATA CCCCTTATCAC  6721 AGGGGACT ATGATGGARA ACCCCTATATA CGATCTACAC GTAAATGACC TATTTTTTGGG  6721 AGGGCACTA TAGATGAAAA GCAGTCTCTC TCTGTGTACA GTAAATGACA TGTTTTTGGG  6721 AGGCACTA ATGATGCAAA GCAGTCTCTC TCTGTGTACA GTAAATGACA TGTTTTTTGGG  6721 AGGCCACA ATGATGCAAA GCAGTCTCTC TCTGTGTACA GTAAATGACA TAGATTTTT  CACCTCGAA TACTACGTTT CCTCAGAGAG AGACACATGT CACCGGGTGG ATAAATTTTT  CACCTCGAA TACTACGTTT CCTCAGAGAG AGACACATGT CACCGGGTGG ATAAATTTTT  TCACGCTCAA ACCAGGACA CTGTGAAACA CTTAACATAA GAAACAAACG CAGGGTCTGG	6121	TATTCTATTT TCACACGTGA	GTTTCTGTTG	CTGGCTGGCT	GACTGGCATT	ATCTATGCTA
6241 TGTACGTACT CACTGGACC CAGACAGAC CATCATTCT CACTGTCTT GAAACAAAGC 6241 TGTACGGTTT GATCGATGAA CGTATTTAAA GCATTTCATG CAATGACAAA GTGCTCAGTA ACATGCCAAA CTAGCTACTT GCATAAATTT CGTAAAGTAC GTTACTGTTT CACGAGGACA 6301 GTGGAAGGCA GGCTGTGACC AGTCTGCCTG CTCCTTACTA TAATTGTGAG GATTTGTTAC CACCTTCCGT CCGACACTGG TCAGACGGAC GAGGAATGAT ATTAACACTC CTAAACAATG ACCTTGCAT GTACCTCCGG ACTGGAACAC CCCCGTGTCC CACCTTGGAA GTCCTTATA ACCTTGTCAT GTACCTCCGG ACTGGAACAC CCCCGTGTCC CACCTTGGAA TCGACTTATA TCACACACAG AGTTCTCCTT CAGTCCCATG ATCGACTCAC GAGTTAGAGG TCCATGATAT TCACACACAG AGTTCTCCTT CAGTCCCATG ATCGACTCAC GAGTTAGAGG TCCATGATAT 6481 TATACATTTG CCCGTTTTAT CTCTAATGTG AAATAAATCC CCAAACACTT GTTTATCGTG ATATGTAAAC GGGCAAAATA GAGATTACAC TTTATTTAGG GGTTTGGAA CAAATAGCAC 6541 TAGCGTACCT AAAAGACTAT TCTATTATGG GTGTCCCCAC TTTCTTGGTT TGGTCACCCC ATCGCATGGA TTTTCTGATA AGATAATACC CACAGGGGTG AAAGAACCAA ACCAGTGGGG 6601 GATCCCCCGG TCTTCTGCTG TATCTAGAAC AGATCATAT TTACTACATA CCCCTTATCAC 6661 TTTCCATATG ATCTGTTGTC TGGAGGATATAT GCTACATGAT TTACTACATA CCCCTTATCAC 6661 TTTCCATATG ATCTGTTGTC TGGAGGATATAT GCTACATGAT TACTACATA CCCCTTATCAC 6721 AGTGCACCT ATGATGCAACA GCACTCTTCT CATTACACA GTAAAATACCA TGTTTTTGGG 6721 AGTGCACCT ATGATGCAACA GCACTCTTC TCTGTTACA GTGCCCCAC TATTTAAAAA CCCTTATCAC 6721 AGTGCACCT ATGATGCAACA GCACTCTTC TCTGTGTACA GTGCCCCAC TATTTAAAAA CCCTTATCAC CTACACTCAAC ATGATGCAACA GCACTCTCT TCTGTGTACA GTGCCCCAC TATTTAAAAA CCCTTATCAC CTACACTCAACA ACCACAGACA ACCTCATATA CAACACACT CAACACACT TCTTTTTGGG 6721 ACTGCACCAA ACCACAGACAA ACCACAGTT CAACACACAC TATTTAAAAA CCCTTACACACAACAACAACAACAACAACAACAACAACACATGT CACGGGGTGG ATAAATTTTT CACACTCCAA ACCACAACAACAACAACAACAACAACAACAACAAC	V121	ATAAGATAAA AGTGTGCACT	CAAAGACAAC	GACCGACCGA	CTGACCGTAA	TAGATACGAT
TCAACTTTAG TCCTCACACG GGTCGTCTCG GGTAGTARGA GTGACAGAAA CTTTGTTTCG 6241 TGTACGGTTT GATCGATGAA CGTATTTAAA GCATTTCATG CAATGACAAA GTGCTCAGTA ACATGCCAAA CTAGCTACTT GCATAAATTT CGTAAAGTAC GTTACTGTTT CACGAGTCAT 6301 GTGGAAGGCA GGCTGTGACC AGTCTGCCTG CTCCTTACTA TAATTGTGAG GATTTGTTAC CACCTTCCGT CCGACACTGG TCAGACGGAC GAGGAATGAT ATTAACACTC CTAAACAATG 6361 TGGAACAGTA CATGGAGGCC TGACCTTGTG GGGGCACAGG GTGGAACCTT AGCTGAATAT ACCTTGTCAT GTACCTCCGG ACTGGAACAC CCCCGTGTCC CACCTTGGAA TCGACTTATA 6421 AGTGTGTGTC TCAAGAGGAA GTCAGGGTAC TAGCTCAGTG CTCAATCTCC AGGTACTATA TCACACACAG AGTTCTCCTT CAGTCCCATG ATCGAGTCAC GAGTTAGAGG TCCATGATAT 6481 TATACATTTG CCCGTTTTAT CTCTAATGTG AAATAAATCC CCAAACACTT GTTTATCGTG ATATGTAAAC GGGCAAAATA GAGATTACAC TTTATTTAGG GGTTTGGAA CAAATAGCAC 6541 TAGCGTACCT AAAAGACTAT TCTATTATGG GTGTCCCAC TTTCTTGGTT TGGTCACCCC ATCGCATGGA TTTTCTGATA AGAATAATACC CACAGGGGTG AAAGAACCAA ACCAGTGGGG 6601 GATCCCCCGG TCTTCTGCTG TATCTAGAACAC AGTGACTATA AATGATGTAT GGGAATAGTG CTAGGGGGCC AGAAGACGAC ATAGATCTTG TCACTGATAT TTACTTACTAT CCCTTATCAC 6661 TTTCCATTATG ATCTGTTGTC TGGAGATATA GCTACATGTT CATTTACTGT ACAAAAACCC AAAGGTATAC TAGACAACAG ACCTCCATATA CGATGTACAA GTAAATTGACA TGTTTTTGGG 6721 AGTGCACCTG ATGATGCAAA GCAGTCTCTC TCTGTGTACAC GTGCCCCACC TATTTAAAAA TCACGTCGAA ACCAGAACAG CCCCCAAGAGAG AGACACATGT CACGGGGTGG ATAAAATTTTT TCACGTCGAA ACCAGAACAC CTGTGAAACAC CTTAACATAA GAAACAAACG CAGCGTCTGG 6721 AGTGCACCAA ACCAGAACAC CTGTGAAACAC CTTAACATAA GAAACAAAACC CAGCGGTTGG 6721 AGTGCACCAA ACCAGAACAC CTGTGAAACAC CTTAACATAA GAAACAAAACG CAGCGTCTGG	6181	ACTTGARATC AGGAGTGTGC	CCAGCAGAGC	CCATCATTCT	CACTGTCTTT	GAAACAAAGC
ACATGCCAAA CTAGCTACTT GCATAAATTT CGTAAAGTAC GTACTCAGTA ACATGCCAAA CTAGCTACTT GCATAAATTT CGTAAAGTAC GTTACTGTTT CACGAGTCAT GGGAAGGCA GGCTGTGACC AGTCTGCCTG CTCCTTACTA TAATTGTGAG GATTTGTTAC CACCTTCCGT CCGACACTGG TCAGACGGAC GAGGAATGAT ATTAACACTC CTAAACAATG CACCTTCCGT CCGACACTGG TCAGACGGAC GAGGAATGAT ATTAACACTC CTAAACAATG  G361 TGGAACAGTA CATGGAGGCC TGACCTTGTG GGGGCACAGG GTGGAACCTT AGCTGAATAT ACCTTGTCAT GTACCTCCGG ACTGGAACAC CCCCGTGTCC CACCTTGGAA TCGACTTATA  6421 AGTGTGTGTC TCAAGAGGAA GTCAGGGTAC TAGCTCAGTG CTCAATCTCC AGGTACTATA TCACACACAG AGTTCTCCTT CAGTCCCATG ATCGAGTCAC GAGTTAGAGG TCCATGATAT  6481 TATACATTTG CCCGTTTTAT CTCTAATGTG AAATAAATCC CCAAACACTT GTTTATCGTG ATATGTAAAC GGGCAAAATA GAGATTACAC TTTATTTAGG GGTTTGTGAA CAAATAGCAC 6541 TAGCGTACCT AAAAGACTAT TCTATTATGG GTGTCCCCCC ATCGCATGGA TTTTCTGATA AGATAATACC CACAGGGGTG AAAAGAACCAA ACCAGTGGGG  6601 GATCCCCCGG TCTTCTGCTG TATCTAGAAC AGTGACTATA AATGATGTAT GGGAATAGTG CTAGGGGGCC AGAAGACGAC ATAGATCTTG TCACTGATAT TTACTACATA CCCTTATCAC 6661 TTTCCATATG ATCTGTTGTC TGGAGTATAT GCTACATGTT CATTTACTGT ACAAAAACCC AAAGGATATAC TAGACAACAG ACCTCATATA CGATGTACAA GTAAATGACA TGTTTTTGGG  6721 AGTGCAGCTG ATGATGCAAA GCAGTCTCTC TCTGTGTACA GTGCCCCACC TATTTAAAAA TCACGTCGAC TACTACCGTTT CGTCAGAAGAG AGACACATGT CACGGGGTGG ATAAATTTTT TCACCTCCAA TACCACAACA CCTGAGAAGAA GAACACATAA GAAACAAAACCC CACGGGTGG ATAAATTTTT TCACCTCCAA TACCACCATTT CCGTCAGAAGAA GAACACATGT CACGGGGTGG ATAAAATTTTT TCACCTCCAA TACCACCAACA CCTGTGAAACAC CTTAACATAA GAAACAACAC CACCGGGTGG ATAAAATTTTT TCACCTCCAA TACCACCAACA CCTGTGAAACAC CTTAACATAA GAAACAACAC CAGCGTCTGG	0101	TCAACTTTAG TCCTCACACG	GGTCGTCTCG	GGTAGTAAGA	GTGACAGAAA	CTTTGTTTCG
GTGGAAGGCA GGCTGTGACC AGTCTGCTG CTCCTTACTA TAATTGTGAG GATTTGTTAC CACCTTCCGT CCGACACTGG TCAGACGGAC GAGGAATGAT ATTAACACTC CTAAACAATG  GAGACAGTA CATGGAGGCC TGACCTTGTG GGGGCACAGG GTGGAACCTT AGCTGAATAT ACCTTGTCAT GTACCTCCGG ACTGGAACAC CCCCGTGTCC CACCTTGGAA TCGACTATA  ACCTTGTCAT GTACCTCCGG ACTGGAACAC CCCCGTGTCC CACCTTGGAA TCGACTATA  ACCTTGTCAT GTACCTCCGG ACTGGAACAC CCCCGTGTCC CACCTTGGAA TCGACTATA  ACTGTGTGTC TCAAGAGGAA GTCAGGGTAC TAGCTCAGTG CTCAATCTCC AGGTACTATA  TCACCACACAG AGTTCTCCTT CAGTCCCATG ATCGAGTCAC GAGTTAGAGG TCCATGATAT  TATACATTTG CCCGTTTTAT CTCTAATGTG AAATAAATCC CCAAACACTT GTTTATCGTG  ATATGTAAAC GGGCAAAATA GAGATTACAC TTTATTTAGG GGTTTGTGAA CAAATAGCAC  ATCGCATGGA TTTTCTGATA AGATAATACC CACAGGGGTG AAAGAACCAA ACCAGTGGGG  GATCCCCCGG TCTTCTGCTG TATCTAGAAC AGTGACTATA AATGATGATA GCACATAGGG  CTAGGGGGCC AGAAGACACA ATAGATCTTG TCACTGATAT TTACTACATA CCCTTATCAC  AAAGGTATAC TAGACAACAG ACCTCATATA GCTACATGTT CATTTACTGT ACAAAAACCC  AAAGGTATAC TAGACAACAG ACCTCATATA CCGATGTACA GTACACCAC TGTTTTTGGG  6721 AGTCCACGAC ATGATGCAAA GCAGTCTCC TCTGTGTACA GTGCCCCACC TATTTAAAAA  TCACGTGGAC TACTACGTTT CGTCAGAGAG AGACACATGT CACGGGGTGG ATAAATTTTT  TCACGTCGAC TACTACGTTT CGTCAGAGAG AGACACATGT CACGGGGTGG ATAAATTTTT  TCACGTCGAC TACTACGTTT CGTCAGAGAG AGACACATGT CACGGGGTGG ATAAATTTTT  TCACGTCGAC TACTACGTTT CGTCAGAGAG CTTAACATAA GAAACAACG CAGCGTCTGG	6241	TOTACGGTTT GATCGATGAA	CGTATTTAAA	<b>GCATTTCATG</b>	CAATGACAAA	GTGCTCAGTA
GTGGAAGGCA GGCTGTGACC AGTCTGCCTG CTCCTTACTA TAATTGTGAG GATTTGTTAC CACCTTCCGT CCGACACTGG TCAGACGGAC GAGGAATGAT ATTAACACTC CTAAACAATG TGAGACGAC GAGGAATGAT ATTAACACTC CTAAACAATG ACCTTGCAT GTACCTCCGG ACTGGAACAC CCCCGTGTCC CACCTTGGAA TCGACTTATA ACCTTGTCAT GTACCTCCGG ACTGGAACAC CCCCGTGTCC CACCTTGGAA TCGACTATAA TCACACACAG AGTTCTCCTT CAGTCCCATG ATCGAGTCAC GAGTTAGAGG TCCATGATAT TCACACACAG AGTTCTCCTT CAGTCCCATG ATCGAGTCAC GAGTTAGAGG TCCATGATAT TAACACTTG CCCGTTTTAT CTCTAATGTG AAATAAATCC CCAAACACTT GTTTATCGTG ATATGGTAACC GGGCAAAATA GAGATTACAC TTTATTTAGG GGTTTGTGAA CAAATAGCAC ACCAGTGGGG AAAGAACCAA ACCAGTGGGG TCTCTGATAA AGATAATACC CACAGGGGTG AAAGAACCAA ACCAGTGGGG CTAGGGGGCC AGAAGACGAC ATAGATCTTG TCACTGATAT AATGATGTAT GGGAATAGTG CCCTTATCAC AAAGGACTAT TTGGTTATCACATA CCCTTATCAC CCCTTATCAC AAAGGATATAC TAGACACAG ACCTCATATA CGATGTACAA GTAAATGACA TGTTTTTGGG AAAAGGTATAC TAGACAACAG ACCTCATATA CGATGTACAA GTAAATGACA TGTTTTTGGG AAAAGGTATAC TAGACAACAG ACCTCATATA CGATGTACAA GTAAATGACA TGTTTTTGGG AAAAGGTATAC TAGACAACAA GCAGTCTCTC TCTGTGTACAA GTAAATGACA TGTTTTTGGG AAAAGGTATAC TAGACAACAA GCAGTCTCTC TCTGTGTACAA GTAAATGACA TGTTTTTTGGG ATAAATTTTT TCACGTCGAC TACTACGTTT CGTCAGAGAG AGACACATGT CACGGGGTGG ATAAATTTTT		ACATGCCAAA CTAGCTACTT	GCATAAATTT	CGTAAAGTAC	GTTACTGTTT	CACGAGTCAT
CACCTTCCGT CCGACACTGG TCAGACGGAC GAGGAATGAT ATTAACACTC CTAAACAATG  1 TGGAACAGTA CATGGAGGCC TGACCTTGTG GGGGCACAGG GTGGAACCTT AGCTGAATAT  1 ACCTTGTCAT GTACCTCCGG ACTGGAACAC CCCCGTGTCC CACCTTGGAA TCGACTATA  1 AGTGTGTGTC TCAAGAGGAA GTCAGGGTAC TAGCTCAGTG CTCAATCTCC AGGTACTATA  1 AGTGTGTGTC TCAAGAGGAA GTCAGGGTAC TAGCTCAGTG CTCAATCTCC AGGTACTATA  1 AGTGTAGAC AGTTCTCCTT CAGTCCCATG ATCGAGTCAC GAGTTAGAGG TCCATGATAT  1 ATACATTTG CCCGTTTTAT CTCTAATGTG AAATAAATCC CCAAACACTT GTTTATCGTG  1 ATACGTAAAC GGGCAAAATA GAGATTACAC TTTATTTAGG GGTTTGTGAA ACAATAGCAC  1 ATACGTACCT AAAAGACTAT TCTATTATGG GTGTCCCCAC TTTCTTGGTT TGGTCACCCC  1 ATCGCATGGA TTTTCTGATA AGATAATACC CACAGGGGTG AAAGAACCAA ACCAGTGGGG  1 CTAGGGGGCC AGAAGACGAC ATAGATCTTG TCACTGATAT TTACTACATA CCCTTATCACC  1 AAAGGTATAC ATCTGTTGTC TGGAGTATAT GCTACATGTT CATTTACTGT ACAAAAACCC  1 AAAGGTATAC TAGACAACAG ACCTCATATA CGATGTACAA GTGACCCAC TATTTAGGG  1 AGTGCAGCTG ATGATGCAAA GCAGTCTCTC TCTGTGTACA GTGCCCCACC TATTTAAAAA  1 TCACGTCGAC TACTACGTTT CGTCAGAGAG AGACACATGT CACGGGGTGG ATAAATTTTT  1 ACCGTCGAC TACTACGTTT CGTCAGAGAG CTTAACATAA GAAACAAACG CAGCGTCTGG	6301	CTCCAACCCA CCCTCTGACC	AGTCTGCCTG	CTCCTTACTA	TAATTGTGAG	GATTTGTTAC
TGGAACAGTA CATGGAGGCC TGACCTTGTG GGGGCACAGG GTGGAACCTT AGCTGAATAT ACCTTGTCAT GTACCTCCGG ACTGGAACAC CCCCGTGTCC CACCTTGGAA TCGACTTATA ACCTTGTCAT GTACCTCCGG ACTGGAACAC CCCCGTGTCC CACCTTGGAA TCGACTTATA TCACACACAG AGTTCTCCTT CAGTCCCATG ATCGAGTCAC GAGTTAGAGG TCCATGATAT TCACACACAG AGTTCTCCTT CAGTCCCATG ATCGAGTCAC GAGTTAGAGG TCCATGATAT TCACACACAG AGTTCTCCTT CAGTCCCATG ATCGAGTCAC GAGTTAGAGG TCCATGATAT ATATGTAAAC GGGCAAAATA GAGATTACAC TTTATTTAGG GGTTTGTGAA CAAATAGCAC TAGCGTACCT AAAAGACTAT TCTATTATGG GTGTCCCCAC TTTCTTGGTT TGGTCACCCC ATCGCATGGA TTTTCTGATA AGATAATACC CACAGGGGTG AAAGAACCAA ACCAGTGGGG CTAGGGGGCC AGAAGACGAC ATAGATCTTG TCACTGATAT TTACTACATA CCCTTATCAC TTTCCATATG ATCTGTTGTC TGGAGTATAT GCTACATGTT CATTTACTGT ACAAAAACCC AAAGGTATAC TAGACAACAG ACCTCATATA CGATGTACA GTGCCCCACC TATTTAGAGG TCACGTCGAC ATGATGCAAA GCAGTCTCTC TCTGTGTACA GTGCCCCACC TATTTAAAAA TCACGTCGAC ATGATGCAAA GCAGTCTCTC TCTGTGTACA GTGCCCCACC TATTTAAAAA TCACGTCGAC TACTACGTTT CGTCAGAGAG AGACACATGT CACGGGGTGG ATAAATTTTT TCACGTCGAC ACCTCAGAACA CCTCAGAGAG AGACACATGT CACGGGGTGG ATAAATTTTT TCACGTCGAC ACCCCAGAACA CTGTGAAACA CTTAACATAA GAAACAACG CAGCGTCTGG	0002	CACCTTCCGT CCGACACTGG	TCAGACGGAC	GAGGAATGAT	ATTAACACTC	CTAAACAATG
ACCTTGTCAT GTACCTCCGG ACTGGAACAC CCCCGTGTCC CACCTTGGAA TCGACTTATA  6421 AGTGTGTGTC TCAAGAGGAA GTCAGGGTAC TAGCTCAGTG CTCAATCTCC AGGTACTATA  TCACACACAG AGTTCTCCTT CAGTCCCATG ATCGAGTCAC GAGTTAGAGG TCCATGATAT  6481 TATACATTTG CCCGTTTTAT CTCTAATGTG AAATAAATCC CCAAACACTT GTTTATCGTG  ATATGTAAAC GGGCAAAATA GAGATTACAC TTTATTTAGG GGTTTGTGAA CAAATAGCAC  6541 TAGCGTACCT AAAAGACTAT TCTATTATGG GTGTCCCCAC TTTCTTGGTT TGGTCACCCC  ATCGCATGGA TTTTCTGATA AGATAATACC CACAGGGGTG AAAGAACCAA ACCAGTGGGG  6601 GATCCCCCGG TCTTCTGCTG TATCTAGAAC AGTGACTATA AATGATGTAT GGGAATAGTG  CTAGGGGGCC AGAAGACGAC ATAGATCTTG TCACTGATAT TTACTACATA CCCTTATCAC  6661 TTTCCATATG ATCTGTTGTC TGGAGTATAT GCTACATGTT CATTTACTGT ACAAAAACCC  AAAGGTATAC TAGACAACAG ACCTCATATA CGATGTACA GTAAATGACA TGTTTTTGGG  6721 AGTGCAGCTG ATGATGCAAA GCAGTCTCTC TCTGTGTACA GTGCCCCACC TATTTAAAAA  TCACGTCGAC TACTACGTTT CGTCAGAGAG AGACACATGT CACGGGGTGG ATAAATTTTT  TCACGTCGAC TACTACGTTT CGTCAGAGAG CTTAACATAA GAAACAAACG CAGCGTCTGG	6361	TOCARCACTA CATGGAGGCC	TGACCTTGTG	GGGGCACAGG	GTGGAACCTT	AGCTGÁATAT
AGTGTGTC TCAAGAGGAA GTCAGGGTAC TAGCTCAGTG CTCAATCTCC AGGTACTATA TCACACACAG AGTTCTCCTT CAGTCCCATG ATCGAGTCAC GAGTTAGAGG TCCATGATAT  6481 TATACATTTG CCCGTTTTAT CTCTAATGTG AAATAAATCC CCAAACACTT GTTTATCGTG ATATGTAAAC GGGCAAAATA GAGATTACAC TTTATTTAGG GGTTTGTGAA CAAATAGCAC  6541 TAGCGTACCT AAAAGACTAT TCTATTATGG GTGTCCCCAC TTTCTTGGTT TGGTCACCCC ATCGCATGGA TTTTCTGATA AGATAATACC CACAGGGGTG AAAGAACCAA ACCAGTGGGG  6601 GATCCCCCGG TCTTCTGCTG TATCTAGAAC AGTGACTATA AATGATGTAT GGGAATAGTG CTAGGGGGCC AGAAGACGAC ATAGATCTTG TCACTGATAT TTACTACATA CCCTTATCAC  6661 TTTCCATATG ATCTGTTGTC TGGAGTATAT GCTACATGTT CATTTACTGT ACAAAAACCC AAAGGTATAC TAGACAACAG ACCTCATATA CGATGTACAA GTAAATGACA TGTTTTTGGG  6721 AGTGCAGCTG ATGATGCAAA GCAGTCTCTC TCTGTGTACA GTGCCCCACC TATTTAAAAAA TCACGTCGAC TACTACGTTT CGTCAGAGAG AGACACATGT CACGGGGTGG ATAAATTTTT  CTACGTCGAC TACTACGTTT CGTCAGAGAG AGACACATGT CACGGGGTGG ATAAATTTTT  TCACGTCGAC TACTACGTTT CGTCAGAGAG AGACACATGT CACGGGGTGG ATAAATTTTT  TCACGTACAA NOCCCAGAACA CTGTGAAACA CTTTAACATAA GAAACAAACG CAGCGTCTGG	0501	ACCTTGTCAT GTACCTCCGG	ACTGGAACAC	CCCCGTGTCC	CACCTTGGAA	TCGACTTATA
TCACACAGA AGTTCTCCTT CAGTCCCATG ATGAGTCAC GAGTTAGAGG TCCATGATAT  6481 TATACATTG CCCGTTTTAT CTCTAATGTG AAATAAATCC CCAAACACTT GTTTATCGTG ATATGTAAAC GGGCAAAATA GAGATTACAC TTTATTTAGG GGTTTGTGAA CAAATAGCAC  6541 TAGCGTACCT AAAAGACTAT TCTATTATGG GTGTCCCCAC TTTCTTGGTT TGGTCACCCC ATCGCATGGA TTTTCTGATA AGATAATACC CACAGGGGTG AAAGAACCAA ACCAGTGGGG  6601 GATCCCCCGG TCTTCTGCTG TATCTAGAAC AGTGACTATA AATGATGTAT GGGAATAGTG CTAGGGGGCC AGAAGACGAC ATAGATCTTG TCACTGATAT TTACTACATA CCCTTATCAC  6661 TTTCCATATG ATCTGTTGTC TGGAGTATAT GCTACATGTT CATTTACTGT ACAAAAAACCC AAAGGTATAC TAGACAACAG ACCTCATATA CGATGTACA GTAAATGACA TGTTTTTGGG  6721 AGTGCAGCTG ATGATGCAAA GCAGTCTCTC TCTGTGTACA GTGCCCCACC TATTTAAAAA TCACGTCGAC TACTACGTTT CGTCAGAGAG AGACACATGT CACGGGGTGG ATAAATTTTT  TCACGTCGAC NCCCAGAACA CTGTGAAACA CTTAACATAA GAAACAAACG CAGCGTCTGG	6421	ACTOTOTO TOADGAGGAA	GTCAGGGTAC	TAGCTCAGTG	CTCAATCTCC	AGGTACTATA
TATACATTIG CCCGTTTAT CTCTAATGTG AAATAAATCC CCAAACACTT GTTTATCGTG ATATGTAAAC GGGCAAAATA GAGATTACAC TTTATTTAGG GGTTTGTGAA CAAATAGCAC  6541 TAGCGTACCT AAAAGACTAT TCTATTATGG GTGTCCCCAC TTTCTTGGTT TGGTCACCCC ATCGCATGGA TTTTCTGATA AGATAATACC CACAGGGGTG AAAGAACCAA ACCAGTGGGG  6601 GATCCCCCGG TCTTCTGCTG TATCTAGAAC AGTGACTATA AATGATGTAT GGGAATAGTG CTAGGGGGCC AGAAGACGAC ATAGATCTTG TCACTGATAT TTACTACATA CCCTTATCAC  6661 TTTCCATATG ATCTGTTGTC TGGAGTATAT GCTACATGTT CATTTACTGT ACAAAAAACCC AAAGGTATAC TAGACAACAG ACCTCATATA CGATGTACA GTAAATGACA TGTTTTTGGG  6721 AGTGCAGCTG ATGATGCAAA GCAGTCTCTC TCTGTGTACA GTGCCCCACC TATTTAAAAA TCACGTCGAC TACTACGTTT CGTCAGAGAG AGACACATGT CACGGGGTGG ATAAATTTTT  TCACGTCGAC TACTACGTTT CGTCAGAGAG AGACACATGT CACGGGGTGG ATAAATTTTT  TCACGTACAA NCCCAGAACA CTGTGAAACA CTTAACATAA GAAACAAACG CAGCGTCTGG	0122	TCACACACAG AGTTCTCCTT	CAGTCCCATG	ATCGAGTCAC	GAGTTAGAGG	TCCATGATAT
ATATGTAAAC GGGCAAAATA GAGATTACAC TTTATTTAGG GGTTTGTGAA CAAATAGCAC  6541 TAGCGTACCT AAAAGACTAT TCTATTATGG GTGTCCCCAC TTTCTTGGTT TGGTCACCCC ATCGCATGGA TTTTCTGATA AGATAATACC CACAGGGGTG AAAGAACCAA ACCAGTGGGG  6601 GATCCCCCGG TCTTCTGCTG TATCTAGAAC AGTGACTATA AATGATGTAT GGGAATAGTG CTAGGGGGCC AGAAGACGAC ATAGATCTTG TCACTGATAT TTACTACATA CCCTTATCAC  6661 TTTCCATATG ATCTGTTGTC TGGAGTATAT GCTACATGTT CATTTACTGT ACAAAAACCC AAAGGTATAC TAGACAACAG ACCTCATATA CGATGTACAA GTAAATGACA TGTTTTTGGG  6721 AGTGCAGCTG ATGATGCAAA GCAGTCTCTC TCTGTGTACA GTGCCCCACC TATTTAAAAA TCACGTCGAC TACTACGTTT CGTCAGAGAG AGACACATGT CACGGGGTGG ATAAATTTTT  6781 TCACGTACAA NCCCAGAACA CTGTGAAACA CTTAACATAA GAAACAAACG CAGCGTCTGG	6481	SANCASTE COCCUTTAN	CTCTAATGTG	AAATAAATCC	CCAAACACTT	GTTTATCGTG
TAGGGTACCT AAAAGACTAT TCTATTATGG GTGTCCCCAC TTTCTTGGTT TGGTCACCCC ATCGCATGGA TTTTCTGATA AGATAATACC CACAGGGGTG AAAGAACCAA ACCAGTGGGG  GATCCCCCGG TCTTCTGCTG TATCTAGAAC AGTGACTATA AATGATGTAT GGGAATAGTG CTAGGGGGCC AGAAGACGAC ATAGATCTTG TCACTGATAT TTACTACATA CCCTTATCAC  6661 TTTCCATATG ATCTGTTGTC TGGAGTATAT GCTACATGTT CATTTACTGT ACAAAAACCC AAAGGTATAC TAGACAACAG ACCTCATATA CGATGTACAA GTAAATGACA TGTTTTTGGG  6721 AGTGCAGCTG ATGATGCAAA GCAGTCTCTC TCTGTGTACA GTGCCCCACC TATTTAAAAA TCACGTCGAC TACTACGTTT CGTCAGAGAG AGACACATGT CACGGGGTGG ATAAATTTTT  TCACGTCGAC NCCCAGAACA CTGTGAAACA CTTAACATAA GAAACAAACG CAGCGTCTGG	0,01	ATATGTAAAC GGGCAAAATA	GAGATTACAC	TTTATTTAGG	GGTTTGTGAA	CAAATAGCAC
ATCGCATGGA TTTTCTGATA AGATAATACC CACAGGGGTG AAAGAACCAA ACCAGTGGGG  bul  GATCCCCCGG TCTTCTGCTG TATCTAGAAC AGTGACTATA AATGATGTAT GGGAATAGTG  CTAGGGGGCC AGAAGACGAC ATAGATCTTG TCACTGATAT TTACTACATA CCCTTATCAC  AAAGGTATAC TAGACAACAG ACCTCATATA CGATGTACAA GTAAATGACA TGTTTTTGGG  AGTGCAGCTG ATGATGCAAA GCAGTCTCTC TCTGTGTACA GTGCCCCACC TATTTAAAAA  TCACGTCGAC TACTACGTTT CGTCAGAGAG AGACACATGT CACGGGGTGG ATAAATTTTT  TCACGTCGAC NCCCAGAACA CTGTGAAACA CTTAACATAA GAAACAAACG CAGCGTCTGG	6541	TACCCTACCT ADDDCACTAT	TCTATTATGG	GTGTCCCCAC	TTTCTTGGTT	TGGTCACCCC
GATCCCCGG TCTTCTGCTG TATCTAGAAC AGTGACTATA AATGATGTAT GGGAATAGTG CTAGGGGGCC AGAAGACGAC ATAGATCTTG TCACTGATAT TTACTACATA CCCTTATCAC  6661 TTTCCATATG ATCTGTTGTC TGGAGTATAT GCTACATGTT CATTTACTGT ACAAAAACCC AAAGGTATAC TAGACAACAG ACCTCATATA CGATGTACAA GTAAATGACA TGTTTTTGGG 6721 AGTGCAGCTG ATGATGCAAA GCAGTCTCTC TCTGTGTACA GTGCCCCACC TATTTAAAAA TCACGTCGAC TACTACGTTT CGTCAGAGAG AGACACATGT CACGGGGTGG ATAAATTTTT  6781 TCACGTACAA NCCCAGAACA CTGTGAAACA CTTAACATAA GAAACAAACG CAGCGTCTGG	VO.12	ATCGCATGGA TTTTCTGATA	AGATAATAÇC	CACAGGGGTG	AAAGAACCAA	ACCAGTGGGG
GATCCCCGG TCTTCTGCTG TATCTAGAAC AGTGACTATA AATGATGTAT GGGAATAGTG CTAGGGGGCC AGAAGACGAC ATAGATCTTG TCACTGATAT TTACTACATA CCCTTATCAC  AAAGGTATAC ATCTGTTGTC TGGAGTATAT GCTACATGTT CATTTACTGT ACAAAAAACCC AAAAGGTATAC TAGACAACAG ACCTCATATA CGATGTACAA GTAAATGACA TGTTTTTGGG  AGTGCAGCTG ATGATGCAAA GCAGTCTCTC TCTGTGTACA GTGCCCCACC TATTTAAAAA  TCACGTCGAC TACTACGTTT CGTCAGAGAG AGACACATGT CACGGGGTGG ATAAATTTTT  TCACGTACAA NCCCAGAACA CTGTGAAACA CTTAACATAA GAAACAAACG CAGCGTCTGG			Xbal		•	
CTAGGGGGCC AGAAGACGAC ATAGATCTTG TCACTGATAT TTACTACATA CCCTTATCAC  6661 TTTCCATATG ATCTGTTGTC TGGAGTATAT GCTACATGTT CATTTACTGT ACAAAAAACCC  AAAGGTATAC TAGACAACAG ACCTCATATA CGATGTACAA GTAAATGACA TGTTTTTGGG  6721 AGTGCAGCTG ATGATGCAAA GCAGTCTCTC TCTGTGTACA GTGCCCCACC TATTTAAAAA  TCACGTCGAC TACTACGTTT CGTCAGAGAG AGACACATGT CACGGGGTGG ATAAATTTTT  CTAL TCACGTACAA NCCCAGAACA CTGTGAAACA CTTAACATAA GAAACAAACG CAGCGTCTGG	6601	GATCCCCGG TCTTCTGCTG	TATCTAGAAC	AGTGACTATA	AATGATGTAT	GGGAATAGTG
TITCCATATG ATCTGTTGTC TGGAGTATAT GCTACATGTT CATTTACTGT ACAAAAACCC  AAAGGTATAC TAGACAACAG ACCTCATATA CGATGTACAA GTAAATGACA TGTTTTTGGG  AGTGCAGCTG ATGATGCAAA GCAGTCTCTC TCTGTGTACA GTGCCCCACC TATTTAAAAA  TCACGTCGAC TACTACGTTT CGTCAGAGAG AGACACATGT CACGGGGTGG ATAAATTTTT  TCACGTACAA NCCCAGAACA CTGTGAAACA CTTAACATAA GAAACAAACG CAGGGTCTGG		CTAGGGGGCC AGAAGACGAC	ATAGATCTTG	TCACTGATAT	TTACTACATA	CCCTTATCAC
AAAGGTATAC TAGACAACAG ACCTCATATA CGATGTACAA GTAAATGACA TGTTTTTGGG 6721 AGTGCAGCTG ATGATGCAAA GCAGTCTCTC TCTGTGTACA GTGCCCCACC TATTTAAAAA TCACGTCGAC TACTACGTTT CGTCAGAGAG AGACACATGT CACGGGGTGG ATAAATTTTT TCACGTACAA NCCCAGAACA CTGTGAAACA CTTAACATAA GAAACAAACG CAGCGTCTGG	6661	TTTCCATATG ATCTGTTGTC	TGGAGTATAT	GCTACATGTT	CATTTACTGT	ACAAAAACCC
6721 AGTGCAGCTG ATGATGCAAA GCAGTCTCTC TCTGTGTACA GTGCCCCACC TATTTAAAAA  TCACGTCGAC TACTACGTTT CGTCAGAGAG AGACACATGT CACGGGGTGG ATAAATTTTT  TCACGTACAA NCCCAGAACA CTGTGAAACA CTTAACATAA GAAACAAACG CAGCGTCTGG		AAAGGTATAC TAGACAACAG	ACCTCATATA	CGATGTACAA	GTAAATGACA	TGTTTTTGGG
TCACGTCGAC TACTACGTTT CGTCAGAGAG AGACACATGT CACGGGGTGG ATACATTTT	6721	ACTGCAGCTG ATGATGCAAA	GCAGTCTCTC'	TCTGTGTACA	GTGCCCCACC	TATTTAAAAA
TORCETACAR NOCCAGARCA CTGTGAAACA CTTAACATAA GAAACAAACG CAGCGTCTGG		TCACGTCGAC TACTACGTTT	CGTCAGAGAG	AGACACATGT	CACGGGGTGG	ATAAATTTTT
AGTGCATGTT NGGGTCTTGT GACACTTTGT GAATTGTATT CTTTGTTTGC GTCGCAGACC	6781	TCACGTACAA NCCCAGAACA	CTGTGAAACA	CTTAACATAA	GAAACAAACG	CAGCGTCTGG
		AGTGCATGTT NGGGTCTTGT	GACACTTTGT	GAATTGTATT	CTTTGTTTGC	GICGCAGACC

6841	ATTCTTTCCA	AGGAGAGCAG	CTTTCTCCAC	AGGAACACAG	TAACAAAAGA	GGTCCGCCGC
	TAAGAAAGGT	TCCTCTCGTC	GAAAGAGGTG	"TCCTTGTGTC	ATTGTTTTCT	CCAGGCGGCG
6901	CATCCACACC	CAGCCAAGAC	ACCTCAGAGG	CCATAGGGAC	AACCTCCTTG	CTGGCCAACA
	GTAGGTGTGG	GTCGGTTCTG	TGGAGTCTCC	GGTATCCCTG	TTGGAGGAAC	GACCGGTTGT
6961	CCTGCTGGAG	CAGGGCACAG	GTCCCAGCAA	CTGATCCTCA	GTGGATGGGT	CCGCAGTCAA
	GGACGACCTC	GTCCCGTGTC	CAGGGTCGTT	GACTAGGAGT	CACCTACCCA	GGCGTCAGTT
					Hindill	EcoRY
7021	AGCCTTAATG	GGCTCTCTTT	TGAAGGGGAA	AGAAANNTTT		ATATCCAACA
	TCGGAATTAC	CCGAGAGAAA	ACTTCCCCTT	TCTTTNNAAA	GTTCGAATAC	TATAGGTTGT
7081	TTATTATAGT	TGATGAGTTA	GTAAATTCCG	AAAAAAAAAG	ATGATTTAT	ATGTATGACA
	AATAATATCA	ACTACTCAAT	CATTTAAGGC	TTTTTTTTC	TACTAAAATA	TACATACTGT .
7141	TAAAAAAAT	CTTTGTAAAG	TGCGCAAGTG	CAATAATTTA	AAGAGGTCTT	ATCTTTGCAT
	ATTTTTTTA	GAAACATTTC	ACGCGTTCAC	GTTATTAAAT	TTCTCCAGAA	TAGAAACGTA
7201	TTATAAATTA	TAAATATTGT	ACATGTGTGT	AATTŢTTCAT	GTATTCATTT	GCAGTCTTTG
	TAATTTATAA	ATTTATAACA	TGTACACACA	TTAAAAAGTA	CATAAGTAAA	CGTCAGAAAC
7261	TATTTAAAAA	AACTTTACTG	TTATGTTTGT	ATANTAGAAC	ATTAATCATT	TATTATAACT
	ATAAATTTTT	TTGAAATGAC	AATACAAACA	TATTATCTTG	TAATTAGTAA	ATAATATTGA
7321	CAGACAAGGT	GTAAATAAAT	TCATAATTCA	AACAGCCAGT	ATATATGCAT	ATATGGGTGT
	GTCTGTTCCA	CATTTATTTA	AGTATTAAGT	TTGTCGGTCA	TATATACGTA	TATACCCACA
7381	TACATTGCAA	AAATCTCTAT	CTTTGTTCTA	TTCACATGCT	TAAAGAAGTA	AGAAATCTTT
	ATGTAACGTT	TTTAGAGATA	GAAACAAGAT	AAGTGTACGA	ATTTCTTCAT	TCTTTAGAAA
7441	TGTGGATATG	TAATTATACA	TATAAAGTAT	ATATATATGT	ATGATACATG	AAATATATTT
	ACACCTATAC	ATTAATATGT	ATATTTCATA	TATATATACA	TACTATGTAC	<b>AAATATATT</b>
7501	AGAAATGTTC	AATTTTAA	TGGATATTCT	TTGGTGTGAA	TAATTGAATA	CAACATTTTT
	TCTTTACAAG	TATTAAAATT	ACCTATAAGA	AACCACACTT	ATTAACTTAT	GTTGTAAAAA
7561	AAAATGAAAA	AAAAAAAAA	C			•
	TTTTACTTTT '	TTTTTTTTT	G <sup>.</sup>	,		

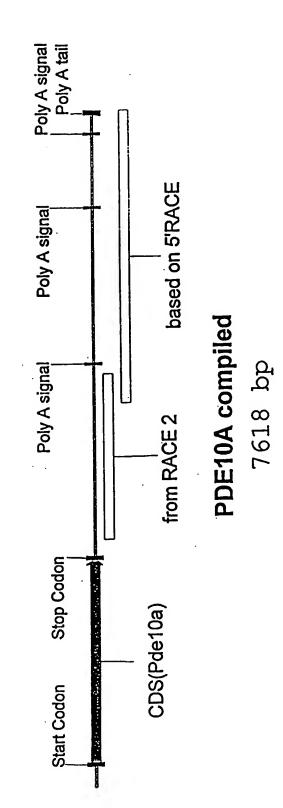
Figure 16



PDE10a and RACEs compiled 7581 bp.

Figure 17

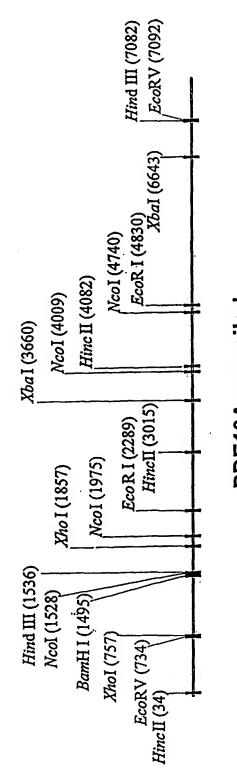
PDE10A compiled - coding sequence and features



34/41

Figure 18

# PDE10A compiled - restriction sites



PDE10A compiled 7618 bp

# Figure 19

## PDE10A compiled

1	CGCCCGGGCA	GGTCTGTTGG	AGGGCAGTTG	GTCAACCTGA	CCAGAGAGAG	CTGAGCTGGA ·
	GCGGGCCCGT	CCAGACAACC	TCCCGTCAAC	CAGTTGGACT	GGTCTCTCTC	GACTCGACCT
61	AGACCCCACT	GATGGTGTGC	TGCCTTTCAG	TCCAGGAAGA	AAGAAAGGAA	GGATTCTGAG
	TCTGGGGTGA	CTACCACACG	ACGGAAAGTC	AGGTCCTTCT	TTCTTTCCTT	CCTAAGACTC.
121	GATTTGGGCA	AAGCCACATT	CCTGGAGAAG	TCTGTATACT	GATGCCAAAC	CCAAGAGCTG
	CTAAACCCGT	TTCGGTGTAA	GGACCTCTTC	AGACATATGA	CTACGGTTTG	GGTTCTCGAC
181	AGCTGCTGAT	GAGGCCCAGG	GAGTAGCCCA	CGCGCCCTGA	GCTGTTGGCT	AGCAAGGCCT
				GCGCGGGACT		
241	TCCTGCTCCA	TGTGGCATGG	AAAAATTATA	TGGTTTGACG	GATGAAAAGG	TGAAGGCCTA
	AGGACGAGGT	ACACCGTACC	TTTTTAATAT	ACCAAACTGC	CTACTTTTCC	ACTTCCGGAT
301	TCTTTCTCTC	CATCCCCAGG	TATTAGATGA	ATTTGTTTCT	GAAAGTGTTA	GTGCAGAGAC
	AGAAAGAGAG	GTAGGGGTCC	ATAATCTACT	TAAACAAAGA	CTTTCACAAT	CACGTCTCTG
361	TGTGGAAAAG	TGGCTGAAGA	GGAAAACCAA	CAAAGCAAAA	GATGAACCAT	CTCCCAAGGA
	ACACCTTTTC	ACCGACTTCT	CCTTTTGGTT	GTTTCGTTTT	CTACTTGGTA	GAGGGTTCCT
421	AGTCAGCAGG	TACCAGGATA	CGAATATGCA	GGGAGTCGTG	TACGAGCTGA	ACAGCTACAT
•	TCAGTCGTCC	ATGGTCCTAT	GCTTATACGT	CCCTCAGCAC	ATGCTCGACT	TGTCGATGTA
481	AGAGCAGCGC	CTGGACACGG	GCGGGGACAA	CCACCTGCTC	CTCTATGAGC	TCAGCAGCAT"
	TCTCGTCGCG	GACCTGTGCC	CGCCCCTGTT	GGTGGACGAG	GAGATACTCG	AGTCGTCGTA
541	CATCAGGATA	GCCACAAAAG	CCGACGGATT	TGCACTGTAC	TTCCTTGGAG	AGTGCAATAA
	GTAGTCCTAT	CGGTGTTTTC	GGCTGCCTAA	ACGTGACATG	AAGGAACCTC	TCACGTTATT
601	TAGCCTGTGT	GTGTTCATAC	CACCCGGGAT	GAAGGAAGGC	CAACCCCGGC	TCATCCCTGC
	ATCGGACACA	CACAAGTATG	GTGGGCCCTA	CTTCCTTCCG	GTTGGGGCCG	AGTAGGGACG
661				TGCCTACGTG		
	TCCCGGGTAG	TGGGTCCCAT	GGTGGTAGAG	ACGGATGCAC	CGGTTCAGAT	CCTTCTGCAA
721				ATTTCCTCGA		
				TAAAGGAGCT		
781	AACCCGCATC	CAGTCTGTTC	TTTGCTTGCC	CATTGTCACT	GCCATTGGAG	ACTTGATTGG ·
				GTAACAGTGA		
841	CATCCTTGAA	CTGTACAGGC	ACTGGGGCAA	AGAGGCCTTC	TGCCTCAGCC	ATCAGGAGGT
				TCTCCGGAAG		
901				AGCAATACAC		
				TCGTTATGTG		
961	TCTCGCCAAA	CAGACCGAAC	TGAATGACTT	CCTACTCGAC	GTATCAAAGA	CATACTTTGA
				GGATGAGCTG		
1021	TAACATAGTT	GCCATAGACT	CTCTACTTGA	ACACATCATG	ATATATGCAA	AAAATCTAGT.
				TGTGTAGTAC		
1081	GAACGCCGAC	CGCTGCGCGC	TCTTCCAGGT	GGACCACAAG	AACAAGGAGC	TGTACTCGGA
				CCTGGTGTTC		
1141	CCTGTTTGAC	ATTGGGGAGG	AGAAGGAGGG	GAAGCCCATC	TTCAAGAAGA	CCAAGGAGAT
				CTTCGGGTAG		
1201	CAGATTTTCC	ATTGAGAAAG	GGATTGCTGG	TCAAGTGGCA	AGAACAGGCG	AAGTCTTGAA
				AGTTCACCGT		
1261	CATTCCCGAT	GCCTACGCGG	ACCCTCGCTT	TAACAGGGAG	GTGGACCTGT	ACACAGGCTA
				ATTGTCCCTC		
1321	CACCACGAGG	AACATTCTGT	GTATGCCCAT	AGTGAGCCGA	CCCTCCCTCT	TTGGCGTGGT
	GTGGTGCTCC	TTGTAAGACA	CATACGGGTA	TOACTOGGCT	CCGICGCACT	AACCGCACCA

### PDEIOA compiled

1381	GCAGATGGTG	AACAAGATCA	GCGGTAGCGC	CTTCTCCAAG	ACAGACGAGA	ACAACTTCAA
		TTGTTCTAGT				
1441	GATGTTTGCT	GTCTTCTGCG	CACTGGCCTT	GCACTGTGCT	AACATGTACC	ACAGGATCCG
	CTACAAACGA	CAGAAGACGC	GTGACCGGAA	CGTGACACGA	TTGTACATGG	TGTCCTAGGC
1501	CCACTCAGAA	TGCATCTACA	GGGTTACCAT	GGAGAAGCTT	TCCTACCACA	GCATCTGCAC
2002	GGTGAGTCTT	ACGTAGATGT	CCCAATGGTA	CCTCTTCGAA	AGGATGGTGT	CGTAGACGTG
1561		TGGCAAGGCC				
1301	GAGGCTCCTC	ACCGTTCCGG	AGTACGCGAA	GTTGGATGGT	CGTGCGTAGA	CGGCCCTGTA
1621		CACTTTGACA				
1021	GCTCGATAAG	GTGAAACTGT	AACCAGGAAA	GCTCTTGTAC	ACCGGACCCT	AGAAACAGAT
1.601		CGGTCTTGTG				
1681	CAIGAICCAI CTACTACCTA	GCCAGAACAC	CCTGTAGGAC	AAAACTTGAA	CTTTTTAACA	CGGCAAAATA
		AAGAAGAACT				
1741	CATGTCTGTG	TTCTTCTTGA	TACCCCCCA	AGGAATGGTG	TTGACCTTCG	TACGTCAGTG
		TGCATGTATG				
1801	GGTGGCACAC	ACGTACATAC	CCMIMCIICA	TTTCTTCTTA	CCGGAGAAGT	GTCTGGAGCT
1861	GCGCAAAGGC	CTGCTAATTG	CGTGTCTGTG	CCTACTCCAC	CTCCTCTCCC	CGAAGTCATT
1921	CAGCTACCTG	CAGAAGTTCG	ACCACCCCT	CCCCCCCC	ATCACCTCCA	CCACCATOGA
		GTCTTCAAGC				
1981	GCAACACCAC	TTCTCCCAGA	CGGTGTCCAT	CCTTCAGCTG	CRECCCCCC	TATACARCAC
		AAGAGGGTCT				
2041	CACCCTGAGC	TCCAGCGAGT	ACGAGCAGGT	GCTGGAGATC	ATCCGCAAAG	CCATCATCGC
	GTGGGACTCG	AGGTCGCTCA	TGCTCGTCCA	CGACCTCTAG	TAGGCGTTTC	GGTAGTAGCG
2101	CACCGACCTC	GCCCTATACT	TTGGGAACAG	GAAGCAGTTG	GAGGAGATGT	ACCAGACAGG
	GTGGCTGGAG	CGGGATATGA	AACCCTTGTC	CTTCGTCAAC	CTCCTCTACA	TGGTCTGTCC
2161	GTCGCTGAAC	CTCCACAACC	AGTCCCATCG	AGACCGTGTC	ATCGGCTTGA	TGATGACTGC
	CAGCGACTTG	GAGGTGTTGG	TCAGGGTAGC	TCTGGCACAG	TAGCCGAACT	ACTACTGACG
2221	CTGTGATCTT	TGCTCTGTGA	CCAAACTATG	GCCAGTTACA	AAATTGACAG	CGAATGATAT
	GACACTAGAA	ACGAGACACT	GGTTTGATAC	CGGTCAATGT	TTTAACTGTC	GCTTACTATA
2281	ATATGCAGAA	TTCTGGGCTG	AGGGTGATGA	GATGAAGAAG	CTGGGCATAC	AGCCCATTCC
	TATACGTCTT	AAGAÇCCGAC	TCCCACTACT	CTACTTCTTC	GACCCGTATG	TCGGGTAAGG
2341	TATGATGGAC	AGAGACAAGC	GAGATGAAGT	CCCTCAAGGG	CAGCTCGGAT	TCTACAATGC
	ATACTACCTG	TCTCTGTTCG	CTCTACTTCA	GGGAGTTCCC	GTCGAGCCTA	AGATGTTACG
2401	mcmccccx mm	CCCTCCTATA	CCACCTTGAC	GCAGATCCTC	CCACCCACAG	AGCCTCTGCT
	ACACCGGTAA	GGGACGATAT	GGTGGAACTG	CGTCTAGGAG	GGTGGGTGTC	TCGGAGACGA
2461	CARGCCTGC	AGGGATAACC	TCAATCAGTG	GGAGAAGGTA	ATTCGCGGGG	AAGAGACAGC
•	CTTCCGGACG	TCCCTATTGG	AGTTAGTCAC	CCTCTTCCAT	TAAGCGCCCC	TTCTCTGTCG
2521	AATCTCCATT	TCAGGCCCAG	GCCCGGCGCC	TAGCAAGAGC	ACACCTGAGA	AGCTGAACGT
2321	TTACACCTAA	AGTCCGGGTC	CGGGCCGCGG	ATCGTTCTCG	TGTGGACTCT	TCGACTTGCA
2581	CANGGTTGAA	GACTGATCCT	GAAGTGACGT	CCTGATGTCT	GCCCAGCAAC	CGACTCAACC
2301	CTTCCAACTT	CTGACTAGGA	CTTCACTGCA	GGACTACAGA	CGGGTCGTTG	GCTGAGTTGG
2641	<b>ずらしずずしずらずら</b>	ACTTCGTTCT	TTTTGTTTC	AAGGGGTGAA	AACCCCCTGT	CAGAAGGTAC
2041	ACGAAGACAC	TGAAGCAAGA	AAAACAAAAG	TTCCCCACTI	TTGGGGGACA	GTCTTCCATG
2201	していこうごうごう	CCATGTGAAG	CAGACGACTC	CCTGCTTGCC	GCACACACCT	CGGACAGTGA
2/01	CCICCCAIAI	GGTACACTTC	GTCTGCTGAG	GGACGAACGG	CGTGTGTGGA	GCCTGTCACT
	GCUGCGIVIV				•	

### PDE10A compiled

2761	1 GUAACCCAGG CTCTGCCGTG TTCAGACGTC GGCTACTCC	G TGGCTCCACC TGACCTCCGA
	CGTTGGGTCC GAGACGGCAC AAGTCTGCAG CCGATGAGG	
2821	1 ATGCTATTTG CTCCCAGGCC AGCACTGCAC TGTCTGGAG	G GGGCAGAGAC CACAGGAGAG
	TACGATAAAC GAGGGTCCGG TCGTGACGTG ACAGACCTC	
2881	1 GTTCTTGCCT GCATCCTCCC ATGAGGGTGT GGCCAGTTC	CTAGTTCTGT GCCATGCTGC
	CAAGAACGGA CGTAGGAGGG TACTCCCACA CCGGTCAAGG	
2941	1 TGCTTGGTGG CATTGGTTAG GAATGGGACA CACGCCCCT	GTTGTGAAGT TTACATGTGA
	ACGAACCACC GTAACCAATC CTTACCCTGT GTGCGGGGAI	CAACACTTCA AATGTACACT
3001	1 CCTTCTTATA GGTTAACTGA GTTTGTGGCC TGGGACACA	GTAATGAAGG TCACAGTCCA
	GGAAGAATAT CCAATTGACT CAAACACCGG ACCCTGTGTA	
3061	1 CAGGTGACAG AGAAATCCAA ACTGTTGATT ACAGGTGCAC	TACAGGTATG CTCTTTCAGT
	GTCCACTGTC TCTTTAGGTT TGACAACTAA TGTCCACGTC	
3121	1 CTATCTGGGG GCACATAGGT GAGTCTGCTC CACTCAGAAC	GAAGCATACC TCTSCCCTCA
• • •	GATAGACCCC CGTGTATCCA CTCAGACGAG GTGAGTCTTC	CTTCGTATGG AGASGGGAGT
3181	1 TCCAGGGGAC ACAGGGTACA TCCCAGGCAT CGGGGAACTO	AAGCTCTCAC TTCAAACCAT
	AGGTCCCCTG TGTCCCATGT AGGGTCCGTA GCCCCTTGAC	TTCGAGAGTG AAGTTTGGTA
3241	1 GTCAAAGAAT TAAAACACCT CCCCTCCCCC TCACTGTAGO	CTTCGGCAAC TGCGCCAATC
	CAGTTTCTTA ATTTTGTGGA GGGGAGGGGG AGTGACATCO	GAAGCCGTTG ACGCGGTTAG
3301	1 CCTTTATACA AAGAAAATAT AAGTAAGGCA TATAAATTTC	CTCCAGCAAG CAAATCTTGT
<u></u>	GGAAATATGT TTCTTTTATA TTCATTCCGT ATATTTAAAG	GAGGTCGTTC GTTTAGAACA
3361		
	CCCATTTTT TTTTTTACAC TTAAAATTGT TGGAGATATA	AAAGTGACAT ACAATACCGT
3421		
	CTTAAAATCA GTGCAGGTTT TGTTTTCTAA TAAGGTCTTC	TATGGAGTAG GATACGGACT
3481		
	TTCGAGGTGT CGTACCGCAG GCAGAGGGTC CCAAGACTAG	T
3541	and the second s	
	TCCGTCCTGT CCTCCTCCAC GTCCCGATGG TGTAACTGGG	
3601		
	TAAGTCTGTA GGTATTCCTT ACGGTTTACG ACATAACTTA	
3661		
	ATCTCTTCGG TCCTGTGGGG ACTCGGAAAG GACCCTTGAG	
3721		
_:	GTGGCACCCC TAAAAGTCCT ATCGTACCTC TGGTCTCTTA	
3781		
2043	AGCCACTCGG AACTCTTCCT TCTCTGACTG GTCTTTGTGA	
3841		
3001	CGTCCTCTTC TATGAAATTC TACTTAGAAA CCCTATCTAA	
3901		
2061		
3961	TTGTAGCGGA GTGGGCTGAA CACTGTAACA CTGTACATGC AACATCGCCT CACCCGACTT GTGACATTGT GACATGTACG	
4021		
4021	AAATGTCACC ATCTCCTCCC CTGCTGTGTC CTACTCCATT TTTACAGTGG TAGAGGAGGG GACGACACAG GATGAGGTAA	
4081		
4001	GTTGTTCTCT TCGATAGTGT TGTGGTCCCG ACACGTGTGC	
	TOTOGICO ACADEMIC ACADEMICA ACA	

### PDE10A compiled

4141	AGCACACAGA TGTATGTAC				
	TCGTGTGTCT ACATACATO	T CGTGTGTGTG	TGTGTGTGTG	GGGTTTTCCT	CTCTTTTCCT
4201	AGAAAACATT TATAAAAAG	C GACAGCTACC	CCCATATTCA	AAAATAGTTC	TTTTCCCTGT
	TCTTTTGTAA ATATTTTTC	G CTGTCGATGG	GGGTATAAGT	TTTTATCAAG	AAAAGGGACA
4261			· · · · · · · · · · · · · · · · · · ·	*** * * * * ** *	1 10 1 10 10 10 1
	TCCCTTTGTC CATCGAGAG				
4321	TTCTCCCAGG GGTGCTCAC			· • • · · · · · · · · · · · · · · · · ·	• • • • • • • • • • • • • • • • • • • •
1561	AAGAGGGTCC CCACGAGTG				
4381	GCAAAAAGCC ATTCGATCC			• • • • • • • • • •	* * *** ** **
4301	CGTTTTTCGG TAAGCTAGG				
4441		*** -***		•	
4441	GGAGGAGCAG CATGTAAGA				
4501	CCTCCTCGTC GTACATTCT				
4501	ATTCTGTCAC CAAGCGTAT				
	TAAGACAGTG GTTCGCATA				• • • • • • • • • • • • • • • • • • • •
4561	TGCTGTAGAA GTAGGGATT				
	ACGACATCTT CATCCCTAA	A ATGTCTTCAG	AGGAACCTAA	ACGGGACGGA	CCCCGTCAAA
4621	TGCAGAGGAA CCTGCCAGA				
•••••	ACGTCTCCTT GGACGGTCT	C TAAATAACCG	ACCAGTCAGA	GAACACTTTA	TCATAGTACA
4681	GAGAAACAGT TTGTAGAAA	A AAACTATACC	TGGGAAGACC	TTTGCAACAT	TGTTCCTTCC
	CTCTTTGTCA AACATCTTT	T TTTGATATGG	ACCCTTCTGG	AAACGTTGTA	ACAAGGAAGG
4741	ATGGGCCAAG ACTCAGTTA	G GAGGCATAAA	TCTGCCCGGA	ATAAACTAGG	CCAGGATACA
	TACCCGGTTC TGAGTCAAT	C CTCCGTATTT	AGACGGGCCT	TATTTGATCC	GGTCCTATGT
4801	GCCATGTTTA GTTAATAAT	T TGGTTTTAGA	ATTCACACAG	GCAGGATTGG	TTTTTTTGTG
	CGGTACAAAT CAATTATTA	A ACCAAAATCT	TAAGTGTGTC	CGTCCTAACC	AAAAAAACAC
4861	TCTTGGCAAG TGGAGCATA	T TTAACATACA	GGCATGGGAA	TCCTGCCTCT	TAGCTTTTCC
	AGAACCGTTC ACCTCGTAT				
4921	CACCCTCTTG TCTCACCAA	G TTTTTTCTCT	CCAAAGGTTT	CCAGGAATTT	CTCATTAATG
	GTGGGAGAAC AGAGTGGTT				
4981	GCTGATGCAA ACTTAGTGA	A TAATAATGAA	TATAAACAAT	GCTCACCTCA	CCAAAATTAT
	CGACTACGTT TGAATCACT				
5041	ATTATTIGCA GTCATTIGT				**
•	TAATAAACGT CAGTAAACA				
5101	GGCCACACAC TGTGGTTAT				
	CCGGTGTGTG ACACCAATA				
5161	TAAGTGCCAA TACCAGTGT				
2202	ATTCACGGTT ATGGTCACA				
5221	TAAACATCAA TTCTATCTC				
JEEI	ATTTGTAGTT AAGATAGAG				
5281	TAAAGTATGC TGGGCTGGT				
, 3201	ATTTCATACG ACCCGACCA				
5241	CTGTCCCCAG CTCCCTCCA				
5341	GACAGGGGTC GAGGGAGGT				
5401	ATATTCTCCC ATAATGGCA				
2401	TATAAGAGGG TATTACCGT				
5461	CCTCTTTTGA GCATGTGTT				
3401	GGAGAAAACT CGTACACAA				
	GONGARAGE COTACACAA	CGIMMANIA	AAAAAAGI		

# Figure 19 (con't) PDE10A compiled

5521	CAAGTGTGTT CATGTAT	STG CTAGATATAI	TAGCACAGCC	TGCCTTCTGC	TGCACAACGC
	GTTCACACAA GTACATA	CAC GATCTATATA	ATCGTGTCGG	ACGGAAGACG	ACGTGTTGCG
5581	CTTAGAGACC CGGCCTT				
	GAATCTCTGG GCCGGAA	GT TACTCGAATO	GAACACGAGA	CAAAGACGAG	AGAATCCAGA
5641	AAACTATGGT GTCAGTT				
	TTTGATACCA CAGTCAAL				
5701	TCGTTTTCAA TGCTGAC		· · · · · · · · · · · · · · · · · · ·		
	AGCAAAAGTT ACGACTG				
5761	TAAGGGACAA CTTTTAAG				
	ATTCCCTGTT GAAAATTC				
5821	TGTCATCACC CCACTTG				
	ACAGTAGTGG GGTGAACT				
5881	GGAACGGTGG CTCCAGGT				
	CCTTGCCACC GAGGTCC				
5941	GTTTTGAAAG TCTCTTC1				
	CAAAACTTTC AGAGAAGA				
6001	CTGTGGCAGT AGGATCTT	~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~		*** ***	
	GACACCGTCA TCCTAGAA				
6061	GTCAGGCTGA CTCGACAG				***************************************
	CAGTCCGACT GAGCTGTC				
6121	CGGCTTCCAC TTGCATGG				
	GCCGAAGGTG AACGTACO				
6181	ACTGGCATTA TCTATGCT			• • •• • •	*****
	TGACCGTAAT AGATACGA				
6241	ACTGTCTTTG AAACAAAG			.,	
	TGACAGAAAC TTTGTTTC				
6301	AATGACAAAG TGCTCAGT		. — — — — —		
•	TTACTGTTTC ACGAGTCA				
6361	AATTGTGAGG ATTTGTTA	CT GGAACAGTAC	ATGGAGGCCT	GACCTTGTGG	GGGCACAGGG
	TTAACACTCC TAAACAAT				
6421	TGGAACCTTA GCTGAATA	TA GTGTGTGTCT	CAAGAGGAAG	TCAGGGTACT	AGCTCAGTGC
	ACCTTGGAAT CGACTTAT	AT CACACACAGA	GTTCTCCTTC	AGTCCCATGA	TCGAGTCACG
6481	TCAATCTCCA GGTACTAT	AT ATACATTTGC	CCGTTTTATC	TCTAATGTGA	AATAAATCCC
	AGTTAGAGGT CCATGATA	TA TATGTAAACG	GGCAAAATAG	AGATTACACT	TTATTTAGGG
6541	CAAACACTTG TTTATCGT	GT AGCGTACCTA	AAAGACTATT	CTATTATGGG	TGTCCCCACT
	GTTTGTGAAC AAATAGCA	CA TCGCATGGAT	TTTCTGATAA	GATAATACCC	ACAGGGGTGA
6601	TTCTTGGTTT GGTCACCC	CG ATCCCCCGGT	CTTCTGCTGT	ATCTAGAACA	GTGACTATAA
	AAGAACCAAA CCAGTGGG	SC TAGGGGGCCA	GAAGACGACA	TAGATCTTGT	CACTGATATT
6661	ATGATGTATG GGAATAGT				
	TACTACATAC CCTTATCA	CA AAGGTATACT	AGACAACAGA	CCTCATATAC	GATGTACAAG
6721	ATTTACTGTA CAAAAACC				
	TAAATGACAT GTTTTTGG		· · · · · · · · · · · · · · · · · · ·	• • • • • • • • • • • • • • • • • • • •	
6781	TGCCCCACCT ATTTAAAA				
	ACGGGGTGGA TAAATTTT			** * **** * * * * * * * * * * * * * * *	
6841	AAACAAACGC AGCGTCTGG				
·	TTTGTTTGCG TCGCAGACO	T AAGAAAGGTT	CCTCTCGTCG	AAAGAGGTGT	CCTTGTGTCA

### PDE10A compiled

6901	AACAAAAGAG	GTCCGCCGCC	ATCCACACCC	AGCCAAGACA	CCTCAGAGGC	CATAGGGACA
	TTGTTTTCTC	CAGGCGGCGG	TAGGTGTGGG	TCGGTTCTGT	GGAGTCTCCG	GTATCCCTGT
6961	ACCTCCTTGC	TGGCCAACAC	CTGCTGGAGC	AGGGCACAGG	TCCCAGCAAC	TGATCCTCAG
	TGGAGGAACG	ACCGGTTGTG	GACGACCTCG	TCCCGTGTCC	AGGGTCGTTG	ACTAGGAGTC
7021	TGGATGGGTC	CGCAGTCAAA	GCCTTAATGG	GCTCTCTTTT	GAAGGGGAAA	GAAANNTTTC
	ACCTACCCAG	GCGTCAGTTT	CGGAATTACC	CGAGAGAAAA	CTTCCCCTTT	CTTTNNAAAG
7081				GATGAGTTAG		
	TTCGAATACT	ATAGGTTGTA	ATAATATCAA	CTACTCAATC	ATTTAAGGCT	TTTTTTTTCT
7141	TGATTTTATA	TGTATGACAT	AAAAAAATC	TTTGTAAAGT	GCGCAAGTGC	AATAATTTAA
	ACTAAAATAT	ACATACTGTA	TTTTTTTAG	AAACATTTCA	CGCGTTCACG	TTATTAAATT
7201				AAATATTGTA		
	TCTCCAGAAT	AGAAACGTAA	ATATŤTAATA	TTTATAACAT	GTACACACAT	TAAAAAGTAC
7261				ACTTTACTGT		
	ATAAGTAAAC	GTCAGAAACA	TAAATTTTTT	TGAAATGACA	ATACAAACAT	ATTATCTTGT
7321				TAAATAAATT		
	AATTAGTAAA					
7381	TATATGCATA					
	ATATACGTAT					
·7441 ·	AAAGAAGTAA					
				TTAATATGTA		
7501	TGATACATGA					
	ACTATGTACT					
7561	AATTGAATAC					
<u></u> .	TTAACTTATG	TTGTAAAAAT	TTTACTTTTT	TTTTTTTTT	TTTTTTTTT	TTTTTTT

### SEQUENCE LISTING

- <110> ROBERTSON, Harold
   DENOVAN-WRIGHT, Eileen
   NOVANEURON, INC.
- <120> GENE NECESSARY FOR STRIATAL FUNCTION, USES THEREOF, AND COMPOUNDS FOR MODULATING SAME
- <130> 36541-0005
- <140>
- <141>
- <150> US60/158,043
- <151> 1999-10-07
- <150> US60/217,765
- <151> 2000-07-12
- <160> 12
- <170> PatentIn Ver. 2.0
- <210> 1
- <211> 3236
- <212> DNA
- <213> mouse
- <400> 1
- cactgaagct ggtccacgtc tataaacagg tgacactggc tgcagcaaaa agccattcga 60 tccacacaaa ttgatcttct atcatcttgg aatctgaatt gcaggagga gcagtatgta 120 agacgaccgt ttaattcagg cattccgaag gcatgagcgc atggattctg tcaccaagcg 180 tataaaagga ccctggcatt gggaaaccta tgacggactg tttttgctgt agaagtaggg 240 attttacaga agtctccttg aatttgccct gcctggggca gttttgcaga ggaacctgcc 300 agagatttat tggctggtca gtctcttgtg aaatagtatc atgtgagaaa cagtttgtag 360 aaaaaaacta tacctgggaa gacctttgca acattgttcc ttccatgggc caagactcag 420 ttaggaggca taaatctgcc cggaataaac taggccagga tacagccatg tttagttaat 480

aatttggttt tagaattcac acaggcagga ttggtttttt tgtgtcttgg caagtggagc 540 atatttaaca tacaggcatg ggaatcctgc ctcttagctt ttcccaccct cttgtctcac 600 caagtttttt ctctccaaag gtttccagga atttctcatt aatggctgat gcaaacttag 660 tgaataataa tgaatataaa caatgctcac ctcaccaaaa ttatattatt tgcagtcatt 720 tgtgataaca caaattttat cgcaatggtt attatttaat ttgtggccac acactgtggt 780 tatcttttgt tgtggttgtt tctgagaaaa tgttcttgga tatgtaagtg ccaataccag 840 tgtgaagtat tgatcccggg cagcaaaata cagcctaagg tttgtaaaca tcaattctat 900 ggtggtggtc agcctcccgc ctgaagagtg accagtgctg gcccgacgga tcgctgagat 1020 atteteceat aatggeaaaa aaataggeag tttgatgtga eetgtttagt gtggetetee 1080 tcttttgagc atgtgttagc atttttattt tatactcatc cagtgaactc tgctcttcca 1140 agtgtgttca tgtatgtgct agatatatta gcacagcctg ccttctgctg cacaacgcct 1200 tagagacccg gcctttcaat gagcttagct tgtgctctgt ttctgctctc ttaggtctaa 1260 actatggtgt cagttttaat agaacaaaag tatgcatctt gccttggctt gagccttttc 1320 gttttcaatg ctgacttctc ccctttctct cctgtgctca ccttaccttt ccagagtgta 1380 agggacaact tttaaggagg cgtgtccctg gtaggggcat ccctgttcac caggtgcctg 1440 tcatcacccc acttgactga catctaccct ggtgactatg ggttcctctt gtttgtaggg 1500 aacggtggct ccaggtggag gcatcaatct gttgggttct ggttcccggc tgcctttggt 1560 tttgaaagtc tcttctctgt atattcctac cctgcatttg ctttgtgtgg tgctgatgct 1620 gtgcgcagta ggattcttgg atgactctcc atcagtcaca gactccccct gttgcaaagt 1680 gtcaggctga ctcgacagtc accgtaaaat ctgagtcagt cacacacagg ctgtcagcca 1740 cggcttccac ttgcatggct attctatttt cacacgtgag tttctgttgc tggctggctg 1800 actggcatta tctatgctaa gttgaaatca ggagtgccca gcagagccca tcattctcac 1860 tgtctttgaa acaaagctgt acggtttgat cgatgaacgt atttaaagca tttcatgcaa 1920 tgacaaagtg ctcagtagtg gaaggcaggc tgtgaccagt ctgcctgctc cttactataa 1980 ttgtgaggat ttgttactgg aacagtacat ggaggcctga ccttgtgggg gcacagggtg 2040 gaaccttagc tgaatatagt gtgtgtctca agaggaagtc agggtactag ctcagtgctc 2100 aatctccagg tactatatat acatttgccc gttttatctc taatgtgaaa taaatcccca 2160 aacacttgtt tatcgtgtag cgtacctaaa agactattct attatgggtg tccccacttt 2220 cttggtttgg tcaccccgat cccccggtct tctgctgtat ctagaacagt gactataaat 2280 gatgtatggg aatagtgttt ccatatgatc tgttgtctgg agtatatgct acatgttcaa 2340 ttactgtaca aaaacccagt gcagctgatg atgcaaagca gtctctctct gtgtacagtg 2400 ccccacctat ttaaaaatca cgtacaascc cagaacactg tgaaacactt aacataagaa 2460 caaacgcagc gtctggattc tttccaagga gagcagcttt ctccacagga acacagtaac 2520 aaaagaggtc cgccgccatc cacacccagc caagacacct cagaggccat agggacaacc 2580 teettgetgg ccaacacetg etggageagg ggeacaggte ccageaactg atectcagtg 2640 gatgggtccg cagtcaaagc cttaatgggc tctcttttga aggggaaaga aagaatttca 2700 agcttatgat atccaacatt attatagttg atgagttagt aaattccaaa aaaaaaagat 2760 gattttatat gtatgacata aaaaaaatct ttgtaaagtg cgcaagtgca ataatttaaa 2820

<210> 2

<211> 5752

<212> DNA

<213> mouse

<400> 2

aagtgtaaat aaaataaaca tctaataaaa aaaattacat accatagagg aacaagataa 60 tttctgccca acttcatacc ctccagcgta tagtgttgag gtttggtctg ttgctgtgta 120 ttgtaatgta atgttaaatt ctctacctga aggtctaggc ctacaagtga attctcatgt 180 ttatagagtt ttgttgtgca aaccttgttc cttaatttaa aactatggtt aaaaaacaaa 240 acaaaactgg ctacagccaa taactgaagg gggttacctt gttgaagggg tggaaaagag 300 acgaggggag atgggccatg agaacttggc caggagaaat agccagtatc tggagtacac 420 cactgaggag gtagccaggc tagcagttag aagagtagat taggggttat ttttccccca 480 ctccacatag ttatcaaagc caaataaaat aaccatagtc tgagtctcat ctatttgtaa 540 gctagttggg tataagatta atttggctgt actacagttt agatttctaa cataggaact 600 atcaaaaact tgctcaaaca agaacatgct gacaatattt taaaatgatt atttatattg 660 tttgcacttt ctaaagtttc ttctaaatgt tccatggtca aattaaaaaa tatacatatt 720 ggctattaaa ttcgtctaag tggggctgga gagatagctc agaggttaag agcactgact 780 gctcttccag aggtcctgag ttcaattccc agcgaccaca tggtggctca cagccatctg 840 taatagatag gatctgacgc cctcttctgg agtgtctgaa gacagctaca atgtactcat 900 atatattaaa taaataatat tagaaaatto ttotaagtgt atcatttata gaatatttaa 960 tatataaagt aaatgcctca ggaaatataa acttggaatt aaatcaaaga acttcatgag 1020 tagtgggcca caaaaaatgt gtaccagggg aagaccggag ggaggggaga aggaagggat 1080 ggagatagaa ttttgcctct gcattccttg ggctggcaca ggtataatgc tgtgggaatt 1140 gggaaactac aaggaagctg caaagctggg cggaactcgt ttccgcaagc tgggctcatc 1200 taagtgtcca tgcatggctg ccacactgca gtgaacttta aaacatttgt gttccagaga 1260 tgtagagatg ctcacaatag tacaaaggcg ggagggaggt atttccagac taagaggaag 1320 aaaaaccatt gctgattaaa catctgcata tgagcgcccc cacctccata cacacaca 1380 cacacacaca cacacaca caaccaaaca gaacaaatac acatgcatgt ctacagcctg 1440 caggaacaaa atggtatgtc tgtgaggaac caggagatgc acaggtccta acctctgtct 1500

cctacaagcc ctgaagtctg gtcagggtca aatgtacaaa agcaggctaa ggaagctgtt 1560 tagtgaaaga tttttttttt caactctagg aacaacctat ttcctaggat ttggagagtg 1620 ctcaggagga aacattcaga caactgatgc tctctgtgta ccccagattc aggtattggg 1680 gtagttagtt gtgctcatgt atgtgctaga tatattagca cagcctgcct tctgctgcac 1740 aacgccttag agacccggcc tttcaatgag cttagcttgt gctctgtttc tgctctctta 1800 ggtctaaact atggtgtcag ttttaataga acaaaagtat gcatcttgcc ttggcttgag 1860 cettttegtt tteaatgetg aetteteece ttteteteet gtgeteaeet tacettteea 1920 gagtgtaagg gacaactttt aaggaggcgt gtccctggta ggggcatccc tgttcaccag 1980 gtgcctgtca tcaccccact tgactgacat ctaccctggt gactatgggt tcctcttgtt 2040 tgtagggaac ggtggctcca ggtggaggca tcaatctgtt gggttctggt tcccggctgc 2100 ctttggtttt gaaagtctct tctctgtata ttcctaccct gcatttgctt tgtgtggtgc 2160 tgatgctgtg cgcagcagga ttcttggatg actctccatc agtcacagac tccccctgtt 2220 gcaaagtgtc aggctgactc gacagtcacc gtaaaatctg agtcagtcac acacaggctg 2280 tcagccacgg cttccacttg catggctatt ctattttcac acgtgagttt ctgttgctgg 2340 ctggctgact ggcattatct atgctaagtt gaaatcaggg gtgcccagca gagcccatca 2400 ttctcactgt ctttgaaaca aagctgtacg gtttgatcga tgaacgtatt taaagcattt 2460 catgcaatga caaagtgctc agtagtggaa ggcaggctgt gaccagtctg cctgctcctt 2520 actataattg tgaggatttg ttactggaac agtacatgga ggcctgacct tgtgggggca 2580 cagggtggaa ccttagctga atatagtgtg tgtctcaaga ggaagtcagg gtactagctc 2640 agtgctcaat ctccaggtac tatatataca tttgcccgtt ttatctctaa tgtgaaataa 2700 ccactttctt ggtttggtca ccccgatccc ccggtcttct gctgtatcta gaacagtgac 2820 tataaatgat gtatgggaat agtgtttcca tatgatctgt tgtctggagt atatgctaca 2880 tgttcattta ctgtacaaaa acccagtgca gctgatgatg caaagcagtc tctctctgtg 2940 tacagtgccc cacctattta aaaatcacgt acttgcccag aacactgtga aacacttaac 3000 ataagaacaa acgcagcgtc tggattcttt ccaaggagag cagctttctc cacaggaaca 3060 cagtaacaaa agaggtccgc cgccatccac acccagccaa gacacctcag aggccatagg 3120 gacaacctcc ttgctggcca acacctgctg gagcaggggc acaggtccca gcaactgatc 3180 ctcagtggat gggtctgcag ccaaagcctt aatgggctct cttttgaagg ggaaagaaag 3240 aatttcaagc ttatgatatc caatattatt atagttgatg agttagtaaa ttccaaaaaa 3300 aaaagatgat tttatatgta tgacataaaa aaaatctttg taaagtgcgc aagtgcaata 3360 atttaaagag gtcttatctt tgcatttata aattataaat attgtacatg tgtgtaattt 3420 ttcatgtatt catttgcagt ctttgtattt aaaaaaactt tactgttatg tttgtataat 3480 agaacattaa tcatttatta taactcagac aaggtgtaaa taaattcata attcaaacag 3540 ccagtatata tgcatatatg ggtgttacat tgcaaaaatc tctatctttg ttctattcac 3600 atgcttaaag aagtaagaaa tcttttgtgg atatgtaatt atacatataa agtatatata 3660 tatgtatgat acatgaaata tatttagaaa tgttcataat tttaatggat attctttggt 3720 aaaatttttt tttttttt ttattccaga gattaaagac actagatctt taaccttgaa 3840

```
gggcaggcaa gaggtcggca atgctgtcaa catagaagtc agggaccatt ttcttcttga 3900
acatgcagtc actttcctga ttgctcttca catcctcaag gctccggaat tccgggggtg 3960
tggtgggctt tgatctcagg actctggagg cagaagcagg cagatctctg tgaatatgag 4020
gccagcctgc actacacaga gctccagacc agtcatggct acatcatgaa accctgtctc 4080
aaaaagaaaa taaaaactgt tgtgtttcta ccatagtgtt aaactcagag tctqaqtaat 4140
gtcgggctga catgctcggg tgtttaacat accttcagct ttgacgaggc gctgaacagt 4200
caaagtctgg ccttggggag cggtggctgt gtttgtgctc aagtccaccg tgaaatcctg 4260
attgtgaatt tggacaaccg tgtccttctt cttggccttc catgcaacct ccaacttcat 4320
gttggtcatt ttgtcaaaac actgtgtgat gtttttatca atatactgcc attccacata 4380
tgtagagatg tagtctgcct ggctttcctt ttctttagcc aatcgaatgc tcttgatcat 4440
gccctcaatc tcatctctag cttttatcac gtctctgcta attcctgaaa cttgaatcga 4500
agttttcttc tggttcatct caatggtgat gttcagttcc ttctgaatct cattcagttt 4560
ctcgtactcc tccatgtcaa agtcactgac acactcatcg tcattggtgt aggaaagctg 4620
ctctttggta atcagttcct ttagccagga gattgttttg ttcacactgt ctacccctga 4680
accacatacc tggaaaactg tgtgctctat tttcttttcc aaaaccaggg tgttcttttt 4740
gggggaaget tgettgggaa agecaagaaa ggetaaagag aaaatggaaa ttaatgttte 4800
ttttactccc ttcaacatca aggttaggaa tatgtatttc ataaaagcta acaactcaca 4860
ggcaatctta gacatcactg actgcttggc aggcgactgc ttggggggag ctggagagcc 4920
ttctctttct ttcatgttgt cgtaaaaaaa ttgcagaata tggggctgga agataacaac 4980
tttaactctc ttcacagcct gcactgattt tttctggaca aattcttcaa tggcatctat 5040
tategetttt getactaegt ttgggteetg ttgageattt cetteaaaaa caaaaaaage 5100
acatttttaa aaagtcaagg ttaagatcca cctgcaaaaa aaagctgcaa tataagcgag 5160
gaattctagt tgtcacagga aataaaaatg tctgttccca ctataatcaa tgtagactga 5220
taatattatg ccagcaaata gttttgaagt cctaggcaca gtgggaggag gttttgttcc 5280
acgctgttca taagccaata ccccagcaaa agaccttaaa ggacaacttq taatttqqqa 5340
cattcacate tgteetette atetgatetg geteccagtg teacteteta acaeggteet 5400
tagagggaca atttatecet geetetgett gatettatge atgtatetgt attettecag 5460
ccatccctgg cgacctgatt tttctaaggc acccaaaact gtaagctact tcttataatc 5520
tataattetg ageatattag ttageetgag cetecaggat atetttette cetataetea 5580
gtccagtttt agctgcccag aaggattcaa agctgatcta cgagtagatc actcctgtct 5640
acagcttgtt ccagatettg ttteteaage eetggaagee ateageeagg taagattgta 5700
aaacaatccc tttctaatca tgggtgtggc ccaaagtgaa tggccggaat tc
                                                                  5752
```

<210> 3

<211> 475

<212> DNA

<213> mouse

<400> 3

```
tgtatgggaa tagtgtttcc atatgatctg ttgtctggag tatatgctac atgttcattt 60
 actgtacaaa aacccagtgc agctgatgat gcaaagcagt ctctctctgt gtacagtgcc 120
 ccacctattt aaaaatcacg tacttgccca gaacactgtg aaacacttaa cataagaaca 180
 aacgcagcgt ctggattctt tccaaggaga gcagctttct ccacaggaac acagtaacaa 240
 aagaggtccg ccgccatcca cacccagcca agacacctca gaggccatag ggacaacctc 300
 cttgctggcc aacacctgct ggagcagggg cacaggtccc agcaactgat cctcagtgga 360
 tgggtctgca gccaaagcct taatgggctc tcttttgaag gggaaagaaa gaatttcaag 420.
 cttatgatat ccaatattat tatagttgat gagttagtaa attccaaaaa aaaaa
                                                                    475
 <210> 4
 <211> 20
 <212> DNA
<213> Artificial Sequence
<220>
<223> Description of Artificial Sequence:primer
<400> 4
agggctgtca atcatgctgg
                                                                   20
<210> 5
<211> 20
<212> DNA
<213> Artificial Sequence
<220>
<223> Description of Artificial Sequence:primer
<400> 5
aaactcacgg tcggtgcagc
                                                                   20
<210> 6
<211> 24
<212> DNA
<213> Artificial Sequence
<220>
<223> Description of Artificial Sequence:probe
```

<400> 6	
attaaccete actaaatget gtat	24
<210> 7	
<211> 30	
<212> DNA	
<213> Artificial Sequence	
<220>	
<223> Description of Artificial Sequence:probe	
<400> 7	
cattatgctg agtgatatct tttttttcg	30
<210> 8	
<211> 38 <212> DNA	
<213> Artificial Sequence	
· · · · · · · · · · · · · · · · · · ·	
<220>	
<223> Description of Artificial Sequence:probe	
<400> 8	
gaacatgtag catatactcc agacaacaga tcatatgg	38
<210> 9	
<211> 32	
<212> DNA	
<213> Artificial Sequence	•
<220>	
<223> Description of Artificial Sequence:probe	
<400> 9	
cagcttctcc acaggaacac agtaacaaag ag	32
ay	34
<210> 10	
<211> 35	•
<212> DNA	

PCT/CA00/01188

WO 01/24781

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence:primer

<400> 10

ctatttcaca agagactgac cagccaataa atctc

35 .

<210> 11

<211> 7581

<212> DNA

<213> Unknown

<400> 11

cgcccgggca ggtctgttgg agggcagttg gtcaacctga ccagagagag ctgagctgga 60 agaccccact gatggtgtgc tgcctttcag tccaggaaga aagaaaggaa ggattctgag 120 gatttgggca aagccacatt cctggagaag tctgtatact gatgccaaac ccaagagctg 180 agctgctgat gaggcccagg gagtagccca cgcgccctga gctgttggct agcaaggcct 240 tectgeteca tgtggcatgg aaaaattata tggtttgacg gatgaaaagg tgaaggceta 300 tetttetete catececagg tattagatga atttgtttet gaaagtgtta gtgcagagae 360 tgtggaaaag tggctgaaga ggaaaaccaa caaagcaaaa gatgaaccat ctcccaagga 420 agtcagcagg taccaggata cgaatatgca gggagtcgtg tacgagctga acagctacat 480 agagcagcgc ctggacacgg gcggggacaa ccacctgctc ctctatgagc tcagcagcat 540 catcaggata gccacaaaag ccgacggatt tgcactgtac ttccttggag agtgcaataa 600 tagcctgtgt gtgttcatac cacccgggat gaaggaaggc caaccccggc tcatccctgc 660 agggcccatc acccagggta ccaccatctc tgcctacgtg gccaagtcta ggaagacgtt 720 gttggtagag gatatcettg gggatgageg attteetega ggtaetggee tggaatcagg 780 aacccgcatc cagtctgttc tttgcttgcc cattgtcact gccattggag acttgattgg 840 catcettgaa etgtacagge aetggggeaa agaggeette tgeeteagee atcaggaggt 900 tgcaacagcc aatcttgctt gggcttccgt agcaatacac caggtgcagg tgtgtagagg 960 tctcgccaaa cagaccgaac tgaatgactt cctactcgac gtatcaaaga catactttga 1020 taacatagtt gccatagact ctctacttga acacatcatg atatatgcaa aaaatctagt 1080 gaacgccgac cgctgcgcgc tcttccaggt ggaccacaag aacaaggagc tgtactcgga 1140 cctgtttgac attggggagg agaaggaggg gaagcccatc ttcaagaaga ccaaggagat 1200 cagattttcc attgagaaag ggattgctgg tcaagtggca agaacaggcg aagtcttgaa 1260 cattcccgat gcctacgcgg accctcgctt taacagggag gtggacctgt acacaggcta 1320 caccacgagg aacattotgt gtatgcccat agtgagccga ggcagcgtga ttggcgtggt 1380 gcagatggtg aacaagatca gcggtagcgc cttctccaag acagacgaga acaacttcaa 1440 gatgtttgct gtcttctgcg cactggcctt gcactgtgct aacatgtacc acaggatccg 1500

ccactcagaa tgcatctaca gggttaccat ggagaagctt tcctaccaca gcatctgcac 1560 ctccgaggag tggcaaggcc tcatgcgctt caacctacca gcacgcatct gccgggacat 1620 cgagctattc cactttgaca ttggtccttt cgagaacatg tggcctggga tctttgtcta 1680 catgatecat eggtettgtg ggacatectg ttttgaactt gaaaaattgt geegttttat 1740 catgtctgtg aagaagaact atcggcgggt tccttaccac aactggaagc atgcagtcac 1800 ggtggcacac tgcatgtatg ccatacttca aaacaacaat ggcctcttca cagacctcga 1860 gcgcaaaggc ctgctaattg cgtgtctgtg ccatgacctg gaccacaggg gcttcagtaa 1920 cagetacetg cagaagtteg accaeceet ggeggegetg tactecacet ecaecatgga 1980 gcaacaccac ttctcccaga cggtgtccat ccttcagctg gaagggcaca atatcttctc 2040 caccetgage tecagegagt acgageaggt getggagate atecgeaaag ceateatege 2100 caccgacctc gccctatact ttgggaacag gaagcagttg gaggagatgt accagacagg 2160 gtcgctgaac ctccacaacc agtcccatcg agaccgtgtc atcggcttga tgatgactgc 2220 ctgtgatctt tgctctgtga ccaaactatg gccagttaca aaattgacag cgaatgatat 2280 atatgcagaa ttctgggctg agggtgatga gatgaagaag ctgggcatac agcccattcc 2340 tatgatggac agagacaagc gagatgaagt ccctcaaggg cagctcggat tctacaatgc 2400 tgtggccatt ccctgctata ccaccttgac gcagatcctc ccacccacag agcctctgct 2460 gaaggcctgc agggataacc tcaatcagtg ggagaaggta attcgcgggg aagagacagc 2520 aatgtggatt teaggeeeag geeeggegee tageaagage acacetgaga agetgaaegt 2580 gaaggttgaa gactgatcct gaagtgacgt cctgatgtct gcccagcaac cgactcaacc 2640 tgcttctgtg acttcgttct ttttgttttc aaggggtgaa aaccccctgt cagaaggtac 2700 cgtcgcatat ccatgtgaag cagacgactc cctgcttgcc gcacacacct cggacagtga 2760 gcaacccagg ctctgccgtg ttcagacgtc ggctactccg tggctccacc tgacctccga 2820 atgetatttg eteccaggee ageaetgeae tgtetggagg gggeagagae caeaggagag 2880 gttcttgcct gcatcctccc atgagggtgt ggccagttcc ctagttctgt gccatgctgc 2940 tgcttggtgg cattggttag gaatgggaca cacgcccctt gttgtgaagt ttacatgtga 3000 ccttcttata ggttaactga gtttgtggcc tggacacatg taatgaaggt cacagtccac 3060 aggtgacaga gaaatccaaa ctgttgatta caggtgcact acaggtatgc tctttcagtc 3120 tatctggggg cacataggtg agtctgctcc actcagaann aagcatacct ctgccctcat 3180 ccaggggaca cagggtacat cccaggcatc ggggaactga agctctcact tcaaaccatg 3240 tcaaagaatt aaaacacctc ccctccccct cactgtagcc ttcgacaact gcgccaatcc 3300 ctttatacaa agaaaataaa agtaaggcat ataaatttcc tccagcaagc aaatcttgtg 3360 ggtaaaaaaa aagcatgtga atnntaacaa cntctanant ntcncngnat gttatggcag 3420 aattttagtc acgtccaaaa caaaaagatt attccagaag atacctcatc ctatgcctqa 3480 aaggeteeac ageatggegt cegteteeca gggttetgat cegteteete aeggtgcaat 3540 caggcaggac agagaggagg gctgcagggc taccacattg acccagaagg tatctcctct 3600 caccattcag acatccataa ggaatgccaa atgctgtatt gaatagttct ctgtgtgact 3660 ttctagagaa gccaggacac cctgagcctt tccnggggaa ctctaaggag tcacaggttc 3720 acaccgtggg gattttcagg atagcatgga gacagagatc cggtcgttgt tctcactcgt 3780 gagcettgag aaggagagae tgaccagaaa cactcactca gcactetgca ggagcaggag 3840

aagatacttt aagatgaatc ttggatagat tttgatacac ccaataccat acacacagga 3900 gcttggcatt tgcaaagtct attcagtttc cttccgcgct ctgacccacg gttgtagcgg 3960 agtgggctga acactgtaac actgtacatg cgatttcccc atgggcttct aaaatgtcac 4020 catctcctcc cctgctgtgt cctactccat ttactggtta caaggtgatg tcaacaagag 4080 aagctatcac aacaccaggg ctgtgcacac gtgcacacac atgtatgcac aagcacacag 4140 atgtatgtac agcacacaca cacacaca ccccaaaagg agagaaaagg aagaaaacat 4200 ttataaaaag cgacagctac cccatatcaa aatagtcttt cctgtaggaa acaggagctc 4260 tccataagga attatcatga gtgtgttctc ccatcagtgc actctcccag gggtgctcac 4320 tgaagctggt ccacrtctat aaacaggtga cactggctgc agcaaaaagc cattcgatcc 4380 acacaaattg atcttctatc atcttggaat ctgaattgca gggaggagca gyatgtaaga 4440 cgaccgttta attcaggcat tccgaaggca tgagcgcatg gattctrtca ccaagcgtat 4500 aaaaggaccc tggcattggg aaacctatga cggactgttt ttgctgtaga agtagggatt 4560 ttacagaagt ctccttgrat ttgccctgcc tggggcagtt ttgcagagga acctgccaga 4620 gatttattgg ctggtcagtc tcttgtgaaa tagtatcatg tgagaaacag tttgtagaaa 4680 aaaactatac ctgggaagac ctttgcaaca ttgttccttc catgggccaa gactcagtta 4740 ggaggcataa atctgcccgg aataaactag gccaggatac agccatgttt agttaataat 4800 ttggttttag aattcacaca ggcaggattg gtttttttgt gtcttggcaa gtggagcata 4860 tttaacatac aggcatggga atcctgcctc ttagcttttc ccaccctctt gtctcaccaa 4920 gttttttctc tccaaaggtt tccaggaatt tctcattaat ggctgatgca aacttagtga 4980 ataataatga atataaacaa tgctcacctc accaaaatta tattatttgc agtcatttgt 5040 gataacacaa attttatcgc aatggttatt atttaatttg tggccacaca ctgtggttat 5100 cttttgttgt ggttgtttct gagaaaatgt tcttggatat gtaagtgcca ataccagtgt' 5160 gaagtattga tecegggeag caaaatacag eetaaggttt gtaaacatea attetatete 5220 agttcatcag agggcctgag aagctgcggg gcagtgtaaa gtaaagtatg ctgggctggt 5280 ggtggtcagc ctccccttgc caagaagaga gcaattgaat cctgtcccca gctccctcca 5340 cgcctgaaga gtgaccagtg ctggcccgac ggatcgctga gatattctcc cataatggca 5400 aaaaaatagg cagtttgatg tgacctgttt agtgtggctc tcctcttttg agcatgtgtt 5460 agcattttta ttttatactc atccagtgaa ctctgctctt ccaagtgtgt tcatgtatgt 5520 gctagatata ttagcacagc ctgccttctg ctgcacaacg ccttagagac ccggcctttc 5580 aatgagetta gettgtgete tgtttetget etettaggte taaactatgg tgteagtttt 5640 aatagaacaa aagtatgcat cttgccttgg cttgagcctt ttcgttttca atgctgactt 5700 ctcccctttc tctcctgtgc tcaccttacc tttccagagt gtaagggaca acttttaagg 5760 aggcgtgtcc ctggtagggg catccctgtt caccaggtgc ctgtcatcac cccacttgac 5820 tgacatctac cetggtgact atgggtteet ettgtttgta gggaacggtg getecaggtg 5880 gaggcatcaa totgttgggt totggttocc ggctgccttt ggttttgaaa gtctcttctc 5940 tgtatattcc taccetgcat ttgctttgtg tggtgctgat gctgtggcag taggatettg 6000 gatgactete cateagteae agacteece tgttgcaaag tgtcaggetg actegacagt 6060 caccgtaaaa tetgagteag teacacacag getgteagee acggetteea ettgeatgge 6120 tattctattt tcacacgtga gtttctgttg ctggctggct gactggcatt atctatgcta 6180

agttgaaatc aggagtgtgc ccagcagagc ccatcattct cactgtcttt gaaacaaagc 6240 tgtacggttt gatcgatgaa cgtatttaaa gcatttcatg caatgacaaa gtgctcagta 6300 gtggaaggca ggctgtgacc agtctgcctg ctccttacta taattgtgag gatttgttac 6360 tggaacagta catggaggcc tgaccttgtg ggggcacagg gtggaacctt agctgaatat 6420 agtgtgtgtc tcaagaggaa gtcagggtac tagctcagtg ctcaatctcc aggtactata 6480 tatacatttg cccgttttat ctctaatgtg aaataaatcc ccaaacactt gtttatcgtg 6540 tagegtacet aaaagactat tetattatgg gtgteeceae tttettggtt tggteaceee 6600 gatcccccgg tcttctgctg tatctagaac agtgactata aatgatgtat gggaatagtg 6660 tttccatatg atctgttgtc tggagtatat gctacatgtt catttactgt acaaaaaccc 6720 agtgcagctg atgatgcaaa gcagtctctc tctgtgtaca gtgccccacc tatttaaaaa 6780 tcacgtacaa ncccagaaca ctgtgaaaca cttaacataa gaaacaaacg cagcgtctgg 6840 attetteca aggagageag etttetecae aggaacaeag taacaaaaga ggteegeege 6900 catccacacc cagccaagac acctcagagg ccatagggac aacctccttg ctggccaaca 6960 cctgctggag cagggcacag gtcccagcaa ctgatcctca gtggatgggt ccgcagtcaa 7020 agcettaatg ggetetettt tgaaggggaa agaaannttt caagettatg atatecaaca 7080 ttattatagt tgatgagtta gtaaattccg aaaaaaaaag atgattttat atgtatgaca 7140 taaaaaaaat ctttgtaaag tgcgcaagtg caataattta aagaggtctt atctttgcat 7200 ttataaatta taaatattgt acatgtgtgt aatttttcat gtattcattt gcagtctttg 7260 tatttaaaaa aactttactg ttatgtttgt ataatagaac attaatcatt tattataact 7320 cagacaaggt gtaaataaat tcataattca aacagccagt atatatgcat atatgggtgt 7380 tacattgcaa aaatctctat ctttgttcta ttcacatgct taaagaagta agaaatcttt 7440 tgtggatatg taattataca tataaagtat atatatatgt atgatacatg aaatatattt 7500 agaaatgttc ataattttaa tggatattct ttggtgtgaa taattgaata caacattttt 7560 aaaatgaaaa aaaaaaaaaa c 7581

<210> 12

<211> 7618

<212> DNA

<213> mouse

<400> 12

egeceggea ggtetgttgg agggeagttg gteaacetga ceagagagg etgagetgga 60 agaececact gatggtgte tgeettteag teeaggaaga aagaaaggaa ggattetgag 120 gatttgggea aagecacatt eetggagaag tetgtataet gatgeeaaae eeaaggeetg 180 agetgetgat gaggeeeagg gagtageeea egegeeetga getgttgget ageaaggeet 240 teetgeteea tgtggeatgg aaaaattata tggtttgaeg gatgaaaagg tgaaggeeta 300 teetteete eateeeeagg tattagatga atttgttet gaaagtgtta gtgeagagae 360 tgtggaaaag tggetgaaga ggaaaaceaa eaaageaaaa gatgaaceat eteeeaagga 420 agteageagg taceaggata egaatatgea gggagtegtg tacgagetga acagetacat 480

agagcagcgc ctggacacgg gcggggacaa ccacctgctc ctctatgagc tcagcagcat 540 catcaggata gccacaaaag ccgacggatt tgcactgtac ttccttggag agtgcaataa 600 tagcctgtgt gtgttcatac cacccgggat gaaggaaggc caaccccggc tcatccctgc 660 agggcccatc acccagggta ccaccatctc tgcctacgtg gccaagtcta ggaagacgtt 720 gttggtagag gatatccttg gggatgagcg atttcctcga ggtactggcc tggaatcagg 780 aacccgcatc cagtctgttc tttgcttgcc cattgtcact gccattggag acttgattgg 840 catcettgaa etgtacagge aetggggeaa agaggeette tgeeteagee atcaggaggt 900 tgcaacagcc aatcttgctt gggcttccgt agcaatacac caggtgcagg tgtgtagagg 960 tctcgccaaa cagaccgaac tgaatgactt cctactcgac gtatcaaaga catactttga 1020 taacatagtt gccatagact ctctacttga acacatcatg atatatgcaa aaaatctagt 1080 gaacgccgac cgctgcgcgc tcttccaggt ggaccacaag aacaaggagc tgtactcgga 1140 cctgtttgac attggggagg agaaggaggg gaagcccatc ttcaagaaga ccaaggagat 1200 cagattttcc attgagaaag ggattgctgg tcaagtggca agaacaggcg aagtcttgaa 1260 cattcccgat gcctacgcgg accctcgctt taacagggag gtggacctgt acacaggcta 1320 caccacgagg aacattctgt gtatgcccat agtgagccga ggcagcgtga ttggcgtggt 1380 gcagatggtg aacaagatca gcggtagcgc cttctccaag acagacgaga acaacttcaa 1440 gatgtttgct gtcttctgcg cactggcctt gcactgtgct aacatgtacc acaggatccg 1500 ccactcagaa tgcatctaca gggttaccat ggagaagett tectaccaca geatetgeae 1560 ctccgaggag tggcaaggcc tcatgcgctt caacctacca gcacgcatct gccgggacat 1620 cgagctattc cactttgaca ttggtccttt cgagaacatg tggcctggga tctttgtcta 1680 catgatecat eggtettgtg ggacateetg ttttgaaett gaaaaattgt geegttttat 1740 catgtctgtg aagaagaact atcggcgggt tccttaccac aactggaagc atgcagtcac 1800 ggtggcacac tgcatgtatg ccatacttca aaacaacaat ggcctcttca cagacctcga 1860 gegeaaagge etgetaattg egtgtetgtg ceatgacetg gaceaeaggg getteagtaa 1920 cagctacetg cagaagtteg accaecect ggeggegetg tactecacet ceaceatgga 1980 gcaacaccac ttctcccaga cggtgtccat ccttcagctg gaagggcaca atatcttctc 2040 caccetgage tecagegagt acgageaggt getggagate ateegeaaag ceateatege 2100 caccgacctc gccctatact ttgggaacag gaagcagttg gaggagatgt accagacagg 2160 gtcgctgaac ctccacaacc agtcccatcg agaccgtgtc atcggcttga tgatgactgc 2220 ctgtgatctt tgctctgtga ccaaactatg gccagttaca aaattgacag cgaatgatat 2280 atatgcagaa ttctgggctg agggtgatga gatgaagaag ctgggcatac agcccattcc 2340 tatgatggac agagacaagc gagatgaagt ccctcaaggg cagctcggat tctacaatgc 2400 tgtggccatt ccctgctata ccaccttgac gcagatcctc ccacccacag agcctctgct 2460 gaaggcctgc agggataacc tcaatcagtg ggagaaggta attcgcgggg aagagacagc 2520 aatgtggatt tcaggcccag gcccggcgcc tagcaagagc acacctgaga agctgaacgt 2580 gaaggttgaa gactgatect gaagtgacgt cetgatgtet geccageaac egaeteaace 2640 tgcttctgtg acttcgttct ttttgttttc aaggggtgaa aaccccctgt cagaaggtac 2700 cgtcgcatat ccatgtgaag cagacgactc cctgcttgcc gcacacacct cggacagtga 2760 gcaacccagg ctctgccgtg ttcagacgtc ggctactccg tggctccacc tgacctccga 2820

atgctatttg ctcccaggcc agcactgcac tgtctggagg gggcagagac cacaggagag 2880 gttcttgcct gcatcctccc atgagggtgt ggccagttcc ctagttctgt gccatgctgc 2940 tgcttggtgg cattggttag gaatgggaca cacgcccctt gttgtgaagt ttacatgtga 3000 ccttcttata ggttaactga gtttgtggcc tgggacacat gtaatgaagg tcacagtcca 3060 caggtgacag agaaatccaa actgttgatt acaggtgcac tacaggtatg ctctttcagt 3120 ctatctgggg gcacataggt gagtctgctc cactcagaag gaagcatacc tctsccctca 3180 tecaggggae acagggtaca teccaggeat eggggaactg aageteteae tteaaaceat 3240 gtcaaagaat taaaacacct cccctccccc tcactgtagc cttcggcaac tgcgccaatc 3300 cetttataca aagaaaatat aagtaaggea tataaattte etecageaag caaatettgt 3360 gggtaaaaaa aaaaaatgtg aattttaaca acctctatat tttcactgta tgttatggca 3420 gaattttagt cacgtccaaa acaaaagatt attccagaag atacctcatc ctatgcctga 3480 aageteeaca geatggegte egteteecag ggttetgate egteteetea eggtgeaate 3540 aggcaggaca ggaggaggtg cagggctacc acattgaccc agatggtatc tcctctcacc 3600 attcagacat ccataaggaa tgccaaatgc tgtattgaat agttctcctg tgtgactttc 3660 tagagaagcc aggacacccc tgagcctttc ctgggaactc ctaaggaagt cacaggttca 3720 caccgtgggg attttcagga tagcatggag accagagaat cccggttcgg ttgttctcac 3780 teggtgagee ttgagaagga agagaetgae cagaaacaet caeteageae tetggeagga 3840 gcaggagaag atactttaag atgaatcttt gggatagatt ttgatacacc caataccata 3900 cacacaggag cttggcattt gcaaagtcta ttcagtttcc ttccacactc tgacccacgg 3960 ttgtagcgga gtgggctgaa cactgtaaca ctgtacatgc gatttcccca tgggcttcta 4020 aaatgtcacc atctcctccc ctgctgtgtc ctactccatt tactggttac aaggtgatgt 4080 caacaagaga agctatcaca acaccagggc tgtgcacacg tgcacacaca tgtatgcaca 4140 agcacacaga tgtatgtaca gcacacacac acacacacac cccaaaagga gagaaaagga 4200 agaaaacatt tataaaaagc gacagctacc cccatattca aaaatagttc ttttccctgt 4260 agggaaacag gtagctctcc ataaggaaat tatcatgagt gtgttctccc atcagtgcac 4320 tteteccagg ggtgeteaet gaagetggte eaegtetata aacaggtgae aetggetgea 4380 gcaaaaagcc attcgatcca cacaaattga tettetatca tettggaate tgaattgcag 4440 ggaggagcag catgtaagac gaccgtttaa ttcaggcatt ccgaaggcat gagcgcatgg 4500 attetgteae caagegtata aaaggaeeet ggeattggga aaeetatgae ggaetgtttt 4560 tgctgtagaa gtagggattt tacagaagtc tccttggatt tgccctgcct ggggcagttt 4620 tgcagaggaa cctgccagag atttattggc tggtcagtct cttgtgaaat agtatcatgt 4680 gagaaacagt ttgtagaaaa aaactatacc tgggaagacc tttgcaacat tgttccttcc 4740 atgggccaag actcagttag gaggcataaa tetgccegga ataaactagg ccaggataca 4800 gccatgttta gttaataatt tggttttaga attcacacag gcaggattgg tttttttgtg 4860 tcttggcaag tggagcatat ttaacataca ggcatgggaa tcctgcctct tagcttttcc 4920 caccetettg teteaceaag tittitetet ecaaaggitt ecaggaatit eteattaatg 4980 gctgatgcaa acttagtgaa taataatgaa tataaacaat gctcacctca ccaaaattat 5040 attatttgca gtcatttgtg ataacacaaa ttttatcgca atggttatta tttaatttgt 5100 ggccacacac tgtggttatc ttttgttgtg gttgtttctg agaaaatgtt cttggatatg 5160

taagtgccaa taccagtgtg aagtattgat cccgggcagc aaaatacagc ctaaggtttg 5220 taaacatcaa ttctatctca gttcatcaga gggcctgaga agctgcgggg cagtgtaaag 5280 taaagtatgc tgggctggtg gtggtcagcc tccccttgcc aagaagagag caattgaatc 5340 ctgtccccag ctccctccac gcctgaagag tgaccagtgc tggcccgacg gatcgctgag 5400 atattetece ataatggeaa aaaaatagge agtttgatgt gaeetgttta gtgtggetet 5460 cctcttttga gcatgtgtta gcatttttat tttatactca tccagtgaac tctgctcttc 5520 caagtgtgtt catgtatgtg ctagatatat tagcacagec tgeettetge tgeacaaege 5580 cttagagace eggeetttea atgagettag ettgtgetet gtttetgete tettaggtet 5640 aaactatggt gtcagtttta atagaacaaa agtatgcatc ttgccttggc ttgagccttt 5700 togttttcaa tgctgacttc tcccctttct ctcctgtgct caccttacct ttccagagtg 5760 taagggacaa cttttaagga ggcgtgtccc tggtaggggc atccctgttc accaggtgcc 5820 tgtcatcacc ccacttgact gacatctacc ctggtgacta tgggttcctc ttgtttgtag 5880 ggaacggtgg ctccaggtgg aggcatcaat ctgttgggtt ctggttcccg gctgcctttg 5940 gttttgaaag tetettetet gtatatteet accetgeatt tgetttgtgt ggtgetgatg 6000 ctgtggcagt aggatcttgg atgactctcc atcagtcaca gactccccct gttgcaaagt 6060 gtcaggctga ctcgacagtc accgtaaaat ctgagtcagt cacacacagg ctgtcagcca 6120 cggcttccac ttgcatggct attctatttt cacacgtgag tttctgttgc tggctggctg 6180 actggcatta tetatgetaa gttgaaatca ggagtgtgee cageagagee cateattete 6240 actgtctttg aaacaaagct gtacggtttg atcgatgaac gtatttaaag catttcatgc 6300 aatgacaaag tgctcagtag tggaaggcag gctgtgacca gtctgcctgc tccttactat 6360 aattgtgagg atttgttact ggaacagtac atggaggcct gaccttgtgg gggcacaggg 6420 tggaacctta gctgaatata gtgtgtgtct caagaggaag tcagggtact agctcagtgc 6480 tcaatctcca ggtactatat atacatttgc ccgttttatc tctaatgtga aataaatccc 6540 caaacacttg tttatcgtgt agcgtaccta aaagactatt ctattatggg tgtccccact 6600 ttcttggttt ggtcaccccg atcccccggt cttctgctgt atctagaaca gtgactataa 6660 atgatgtatg ggaatagtgt ttccatatga tctgttgtct ggagtatatg ctacatgttc 6720 atttactgta caaaaaccca gtgcagctga tgatgcaaag cagtctctct ctgtgtacag 6780 tgccccacct atttaaaaat cacgtacaan cccagaacac tgtgaaacac ttaacataag 6840 aaacaaacgc agcgtctgga ttctttccaa ggagagcagc tttctccaca ggaacacagt 6900 aacaaaagag gtccgccgcc atccacaccc agccaagaca cctcagaggc catagggaca 6960 acctecttge tggccaacae etgetggage agggcaeagg teccageaae tgatecteag 7020 tggatgggtc cgcagtcaaa gccttaatgg gctctctttt gaaggggaaa gaaanntttc 7080 aagettatga tatecaacat tattatagtt gatgagttag taaatteega aaaaaaaaga 7140 tgattttata tgtatgacat aaaaaaaatc tttgtaaagt gcgcaagtgc aataatttaa 7200 agaggtetta tetttgeatt tataaattat aaatattgta catgtgtgta attttteatg 7260 tattcatttg cagtctttgt atttaaaaaa actttactgt tatgtttgta taatagaaca 7320 ttaatcattt attataactc agacaaggtg taaataaatt cataattcaa acagccagta 7380 tatatgcata tatgggtgtt acattgcaaa aatctctatc tttgttctat tcacatgctt 7440 aaagaagtaa gaaatctttt gtggatatgt aattatacat ataaagtata tatatatgta 7500

tgatacatga aatatatta gaaatgttca taattttaat ggatattctt tggtgtgaat 7560 aattgaatac aacattttta aaatgaaaaa aaaaaaaaa aaaaaaaaa 7618

### (19) World Intellectual Property Organization International Bureau



### 

### (43) International Publication Date 12 April 2001 (12.04.2001)

### PCT

### (10) International Publication Number WO 01/24781. A3

- (51) International Patent Classification7: A61K 31/505, 31/4174, 31/65, 31/4985, 31/522, 31/4745, 31/519, 31/7048, 31/395, A61P 25/14, G01N 33/68
- (74) Agent: HELLER, David, J.; Ridout & Maybee, One Queen Street East, Suite 2400, Toronto, Ontario M5C 3B1 (CA).
- (21) International Application Number: PCT/CA00/01188
- (81) Designated States (national): AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CR, CU, CZ, DE, DK, DM, DZ, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, NO, NZ,
- (22) International Filing Date: 6 October 2000 (06.10.2000)
- PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZW.

(84) Designated States (regional): ARIPO patent (GH, GM,

KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZW), Eurasian

patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European

patent (AT, BE, CH, CY, DE, DK, ES, Fl, FR, GB, GR, IE,

IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

(25) Filing Language:

English

(26) Publication Language:

English

- (30) Priority Data:
  - 2,285,690 7 October 1999 (07.10.1999) CA 60/158,043 7 October 1999 (07.10.1999) US 12 July 2000 (12.07.2000) 60/217,765 US
- (71) Applicant (for all designated States except US): NOVA-NEURON INC. [CA/CA]; 5859 University Avenue, Sir Charles Tupper Medical Bldg., Room 15D7, Halifax, Nova Scotia B3H 4H7 (CA).
- Published:
- with international search report

(72) Inventors; and

- (75) Inventors/Applicants (for US only): ROBERTSON, Harold, A. [CA/CA]; 2384 Clifton Street, Halifax, Nova Scotia B3K 4V1 (CA). DENOVAN-WRIGHT, Eileen, M. [CA/CA]: 22 Fleming Drive, Halifax, Nova Scotia B3P 1A9 (CA).
- (88) Date of publication of the international search report: 7 February 2002

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.



(54) Title: GENE NECESSARY FOR STRIATAL FUNCTION, USES THEREOF, AND COMPOUNDS FOR MODULATING SAME

(57) Abstract: PDE10A, a gene that is normally highly expressed in mammalian striatum and elsewhere, has been found to decrease in expression during the development of CAG repeat disorders such as Huntington's disease. The invention teaches a method for detecting the presence of or the predisposition for a CAG repeat disorder. Compounds which modulate CAG repeat disorders and their uses are taught. Methods for screening for further compounds to modulate CAG repeat disorders are also taught.

International Application No PCT/CA 00/01188

A. CLASSIFICATION OF SUBJECT MATTER IPC 7 A61K31/505 A61K31/4174 A61K31/4985 A61K31/65 A61K31/522 A61K31/7048 A61K31/519 A61K31/4745 A61K31/395 A61P25/14 G01N33/68

According to International Patent Classification (IPC) or to both national classification and IPC

### **B. FIELDS SEARCHED**

Minimum documentation searched (classification system followed by classification symbols)  $IPC\ 7\ A61K\ G01N$ 

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)

C. DOCUME	INTS CONSIDERED TO BE RELEVANT	
Category °	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	US 5 789 395 A (ABRAMSON STEVEN B ET AL) 4 August 1998 (1998-08-04) abstract column 4, line 20 - line 62 column 6, line 59 - line 67 column 7, line 13 - line 30 column 8, line 39 - line 45	1,2,5-7
X	WO 99 48522 A (INOUE JUN ;SAKAMOTO YUJI (JP); YOSHIDA YUKUO (JP); SENJU PHARMA CO) 30 September 1999 (1999-09-30) page 1, paragraph 2 -page 2, paragraph 1	1,4-7
X	US 5 849 290 A (BROWN ROBERT ET AL) 15 December 1998 (1998-12-15) abstract column 6, line 17 - line 51	1,2,5-7

	-/
X Further documents are listed in the continuation of box C.	X Patent family members are listed in annex.
"A" document defining the general state of the art which is not considered to be of particular relevance  "E" earlier document but published on or after the international filing date  "L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)  "O" document referring to an oral disclosure, use, exhibition or other means  "P" document published prior to the international filing date but later than the priority date claimed	"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention  "X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone  "Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.  "&" document member of the same patent family
Date of the actual completion of the international search  27 April 2001	Date of mailing of the international search report  16   10   0
Name and mailing address of the ISA  European Patent Office, P.B. 5818 Patentlaan 2  NL - 2280 HV Rijswijk  Tel. (+31-70) 340-2040, Tx. 31 651 epo nl,  Fax: (+31-70) 340-3016	Authorized officer Cielen, E

		PCT/CA 00/01188
C.(Continu	ation) DOCUMENTS CONSIDERED TO BE RELEVANT	
Category "	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	WO 99 06545 A (MAX PLANCK GESELLSCHAFT; LEHRACH HANS (DE); WANKER ERICH (DE); SCH) 11 February 1999 (1999-02-11) page 1, paragraph 1 - paragraph 4 page 17, paragraph 3 - paragraph 6 claims 20-22	1,2,5-7
X	US 5 750 376 A (REYNOLDS BRENT ET AL) 12 May 1998 (1998-05-12) column 1, line 40 - line 57 column 11, line 26 - line 29 column 18, line 6 - line 9 column 20, line 41 - line 56 column 26, line 63 -column 27, line 32 column 27, line 62 -column 28, line 6	1,4-7
X	EP 0 930 069 A (JOHNSON & JOHNSON MEDICAL LTD) 21 July 1999 (1999-07-21) abstract page 2, line 54 - line 57 example 1 claims 1,6	1,4
X	US 4 367 217 A (GRUBER PETER ET AL) 4 January 1983 (1983-01-04) abstract column 1, line 6 - line 11 example I claim 1	1,4
X	US 5 461 068 A (THALER IRWIN ET AL) 24 October 1995 (1995-10-24) abstract column 5, line 4 - line 14 claims 1,6	1,2
X	US 4 292 312 A (GRIFFON HENRI) 29 September 1981 (1981-09-29) abstract examples claims 9,10	1,2
X	WO 97 03985 A (GLAXO WELLCOME LAB SA; DAUGAN ALAIN CLAUDE MARIE (FR); GELLIBERT F) 6 February 1997 (1997-02-06) abstract page 1, line 3 - line 14 page 2, line 13 -page 3, line 2 page 5, line 15 - line 17 page 12, line 6 - line 12 claims 5,11	1,3
İ		

		PCT/CA 00/01188
	ation) DOCUMENTS CONSIDERED TO BE RELEVANT	
Category °	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	VEMULAPALLI SUBBARAO ET AL: "Antiplatelet and antiproliferative effects of SCH 51866, a novel type 1 and type 5 phosphodiesterase inhibitor." JOURNAL OF CARDIOVASCULAR PHARMACOLOGY, vol. 28, no. 6, 1996, pages 862-869, XP000998260 ISSN: 0160-2446 abstract figures 2-5 page 867, column 2, paragraph 4	1,4
X	ICHIMURA MICHIO ET AL: "KS-505a, an isoform-selective inhibitor of calmodulin-dependent cyclic nucleotide phosphodiesterase." BIOCHEMICAL JOURNAL, vol. 316, no. 1, 1996, pages 311-316, XP000998314 ISSN: 0264-6021 abstract page 312, column 1, paragraph 1 page 315, column 2, paragraph 2 - paragraph 3	1,4
A	US 5 932 423 A (SEILHAMER JEFFREY J ET AL) 3 August 1999 (1999-08-03) abstract column 3, line 8 - line 18 column 4, line 27 - line 45 column 11, line 1 - line 7 column 12, line 24 - line 30 column 12, line 53 - line 63 column 19, line 55 - line 62 column 21, line 53 - line 65 column 22, line 10 - line 14 column 25, line 5 - line 17	1,4-7
A	CEPEDA CARLOS ET AL: "Dopaminergic modulation of early signs of excitotoxicity in visualized rat neostriatal neurons."  EUROPEAN JOURNAL OF NEUROSCIENCE, vol. 10, no. 11, November 1998 (1998-11), pages 3491-3497, XP000998317  ISSN: 0953-816X abstract page 3491, column 1, paragraph 1 figure 3C page 3494, column 2, paragraph 2 page 3495, column 1, paragraph 2 page 3496, column 1, paragraph 3	1,2,5-7

		PC1/CA 00/01188
C.(Continua	ntion) DOCUMENTS CONSIDERED TO BE RELEVANT	
Category °	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	WO 99 42596 A (ICOS CORP) 26 August 1999 (1999-08-26) cited in the application abstract page 2, line 31 -page 3, line 8 page 4, line 27 -page 5, line 19 page 14, line 22 -page 15, line 23 page 18, line 25 - line 30 page 19, line 14 - line 16	
A	FUJISHIGE K ET AL: "CLONING AND CHARACTERIZATION OF A NOVEL HUMAN PHOSPHODIESTERASE THAT HYDROLYZES BOTH CAMP AND CGMP (PDE10A)" JOURNAL OF BIOLOGICAL CHEMISTRY, AMERICAN SOCIETY OF BIOLOGICAL CHEMISTS, BALTIMORE, MD, US, vol. 274, no. 26, 25 June 1999 (1999-06-25), pages 18438-18445, XP000983162 ISSN: 0021-9258 abstract table II page 18443, column 2, paragraph 3 -page 18444, column 1, paragraph 1 page 18444, column 2, paragraph 2	1,4
A	LOUGHNEY K., SNYDER P.B. ET AL.: "Isolation and characterisation of PDE10A, a novel human 3',5'-cyclic nucleotide phosphodiesterase." GENE, vol. 234, 1999, pages 109-117, XP002166002 abstract page 114, column 2, paragraph 3 table 1	1,4
P,X	WO 00 27422 A (BIOGEN INC; GALDES ALPHONSE (US); ONTOGENY INC (US); MAHANTHAPPA N) 18 May 2000 (2000-05-18) page 3, line 25 - line 30 page 4, line 19 -page 7, line 7 page 10, line 6 - line 16 page 11, line 14 - line 18 page 13, line 1 - line 6 page 15, line 15 - line 25	1-4,6

	. DOCUMENTO CONCIDENCE TO DE DEL FUANT	PCT/CA 00/01188
	ation) DOCUMENTS CONSIDERED TO BE RELEVANT  Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Category °	Chanon of document, with indication, where appropriate, or the relevant passages	neievant to ciaim no.
P,X	CHEN MINGHUA ET AL: "Minocycline inhibits caspase-1 and caspase-3 expression and delays mortality in a transgenic mouse model of Huntington disease."  NATURE MEDICINE, vol. 6, no. 7, July 2000 (2000-07), pages 797-801, XP000998293  ISSN: 1078-8956 abstract page 797, column 2, paragraph 2 -page 798, column 1, paragraph 1 figure 1 page 800, column 2, paragraph 2 - paragraph 3	1,2,5-7

International application No. PCT/CA 00/01188

### INTERNATIONAL SEARCH REPORT

Box I	Observations where certain claims were found unsearchable (Continuation of item 1 of first sheet)
This Inte	ernational Search Report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:
1.	Claims Nos.: because they relate to subject matter not required to be searched by this Authority, namely:
2.	Claims Nos.: because they relate to parts of the International Application that do not comply with the prescribed requirements to such an extent that no meaningful International Search can be carried out, specifically:
з	Claims Nos.: because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).
Box II	Observations where unity of invention is lacking (Continuation of item 2 of first sheet)
This Inte	Secretary Authority found multiple inventions in this international application, as follows:
1.	As all required additional search fees were timely paid by the applicant, this International Search Report covers all searchable claims.
2.	As all searchable claims could be searched without effort justifying an additional fee, this Authority did not invite payment of any additional fee.
3.	As only some of the required additional search fees were timely paid by the applicant, this International Search Report covers only those claims for which fees were paid, specifically claims Nos.:
4. X	No required additional search fees were timely paid by the applicant. Consequently, this International Search Report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:  1-7
Remark	on Protest  The additional search fees were accompanied by the applicant's protest.  No protest accompanied the payment of additional search fees.

### FURTHER INFORMATION CONTINUED FROM PCT/ISA/ 210

### 1. Claims: 1-7

A composition and its use for treating a CAG repeat disorder comprising a compound which modulates PDE10A expression and a pharmaceutically acceptable carrier.

### 2. Claims: 8-14

A method for identifying a compound which inhibits or promotes a CAG repeat disorder.

### 3. Claims: 15-19

A method for detecting the presence of or the predisposition for a CAG repeat disorder.

Information on patent family members

Patent document	Publication date	Patent family	Publication
cited in search report	<u> </u>	member(s)	date
US 5789395 A	04-08-1998	AU 718234 B AU 4080897 A	13-04-2000 19-03-1998
		EP 0966525 A	29-12-1999
		US 5919775 A	06-07-1999
		WO 9808480 A	05-03-1998
WO 9948522 A	30-09-1999	AU 2852699 A	18-10-1999
		BR 9908948 A	21-11-2000
		CN 1301173 T	27-06-2001
		EP 1064021 A	03-01-2001
		HU 0101428 A NO 20004679 A	28-09-2001
		PL 342970 A	26-10-2000 16-07-2001
US 5849290 A	15-12-1998	US 5843641 A	01-12-1998
03 3043230 11	10 12 1000	CA 2157041 A	01-09-1994
		EP 0686203 A	13-12-1995
		JP 8510377 T	05-11-1996
4		WO 9419493 A	01-09-1994
WO 9906545 A	11-02-1999	EP 1001987 A	24-05-2000
US 5750376 A	12-05-1998	US 6071889 A	06-06-2000
		US 5981165 A	09-11-1999
		US 6294346 B US 5980885 A	25-09-2001 09-11-1999
		US 5851832 A	22-12-1998
		AU 697894 B	22-10-1998
		AU 8056194 A	29-05-1995
		CA 2175992 A	18-05-1995
		WO 9513364 A CN 1141058 A	18-05-1995 22-01-1997
		EP 0728194 A	28-08-1996
		FI 961855 A	04-06-1996
		JP 9507747 T	12-08-1997
		NO 961859 A	03-07-1996
		AU 714837 B AU 3515295 A	13-01-2000 09-04-1996
		CA 2200709 A	28-03-1996
		WO 9609543 A	28-03-1996
		EP 1130394 A	05-09-2001
		EP 0783693 A	16-07-1997
		FI 971168 A JP 10505754 T	20-03-1997
		JP 10505754 T NO 971245 A	09-06-1998 18-03-1997
		AU 716811 B	09-03-2000
		AU 3836795 A	06-06-1996
		WO 9615226 A	23-05-1996
		CN 1170435 A	14-01-1998
		EP 0792350 A FI 971956 A	03-09-1997 04-07-1997
		JP 10509592 T	22-09-1998
		NO 972171 A	07-07-1997
		AU 665012 B	14-12-1995
		AU 2242592 A	11-02-1993
		CA 2113118 A	21-01-1993
		WO 9301275 A EP 0594669 A	21-01-1993 04-05-1994
		LI 0334003 M	04-00-1994

Information on patent family members

Patent document cited in search report		Publication date		atent family nember(s)		Publication date
US 5750376	Α	<u> </u>	FI	935929	A	02-02-199
00 0,000.0			JP		T	20-10-199
			NO	940056	À	03-03-199
			WO	9416718		04-08-199
			AU	683023		30-10-199
			AU	5147493		09-05-199
			CA	2147162		28-04-199
			WO		Ä	28-04-199
			ËP	0664832		02-08-199
			FI	951677		07-04-199
			JP		T	12-03-199
			NO	951378		07-04-199
			AU	703729		01-04-199
		21 07 1000		0222041		14 07 1004
EP 0930069	Α	21-07-1999	GB	2333041		14-07-1999
			AU	1129699		05-08-1999
			BR	9900046		02-05-2000
			JP	11279080		12-10-1999
			SG	77214	A	19-12-2000
us 4367217	A	04-01-1983	DE	3000979	Α	23-07-1981
			AR	221983		31-03-1981
			AT	6585		15-03-1984
			AU	539618	В	11-10-1984
			AU	6615681	Α	16-07-1981
			CA	1148084	Α	14-06-1983
			CS	8100232	Α	18-06-1984
			DD	156943	Α	06-10-1982
		•	DE	3067058		19-04-1984
			DK	9481		13-07-1981
			EP	0032562		29-07-1981
			ES	498393		16-09-1982
			ES	8207428		16-12-1982
			FI	810058	Α,Β,	13-07-1981
			GB	2067073		22-07-1981
			GR	73102		02-02-1984
			HK	62184		17-08-1984
		•	ĤΩ	184366		28-08-1984
			ΙE	50807		23-07-1986
			IL	61887		30-09-1984
			IN	155158		05-01-1985
			JP	1350517		28-11-1986
			JP	56103111		18-08-1981
			JP	61016368		30-04-1986
			KE	3402		08-06-1984
			MY	55985		31-12-1985
			NO	810061		13-07-1981
			NZ	195994		27-04-1984
			OA	6720		30-06-1982
			PH	17063		24-05-1984
			PL	229157		18-09-1981
			PT	72330		01-02-1981
			SG	28184		15-02-1985
			YU	3981		30-09-1983
			ZA	8100161	Α	29-09-1982
US 5461068	A	24-10-1995	NONE			
	~	ムマーエローエフフリ	NUNE			

Information on patent family members

Patent document cited in search report				Publication date	
		Publication date	Pater mer		
US 4292312	Α .	29-09-1981	DE DK EP	2459046 A 3071091 D 303679 A 0021953 A 2489148 A	09-01-1981 24-10-1985 16-12-1980 07-01-1981 05-03-1982
WO 9703985	A	06-02-1997	BR CA CN CZ EP HR HU JP 1 NO PL SK US	702324 B 6419296 A 9609780 A 2226761 A 1195349 A,B 9800032 A 0846118 A 960321 A 9900006 A 1509535 T 980154 A 324527 A 3898 A 5981527 A	18-02-1999 18-02-1997 09-03-1999 06-02-1997 07-10-1998 17-06-1998 10-06-1998 31-08-1998 28-04-1999 24-08-1999 10-03-1998 08-06-1998 04-11-1998 09-11-1999 07-11-2000
US 5932423	A	03-08-1999	US AU AU CA EP	5798246 A 6080548 A 724247 B 2542797 A 2248675 A 9907742 A	25-08-1998 27-06-2000 14-09-2000 17-10-1997 02-10-1997 14-04-1999 02-10-1997
WO 9942596	A	26-08-1999		3307999 A 1060251 A	06-09-1999 20-12-2000
WO 0027422	Α	18-05-2000		l611600 A l126865 A	29-05-2000 29-08-2001

# This Page is Inserted by IFW Indexing and Scanning Operations and is not part of the Official Record

### **BEST AVAILABLE IMAGES**

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images include but are not limited to the items checked:

BLACK BORDERS

BLACK BURDERS
☐ IMAGE CUT OFF AT TOP, BOTTOM OR SIDES
FADED TEXT OR DRAWING
BLURRED OR ILLEGIBLE TEXT OR DRAWING
☐ SKEWED/SLANTED IMAGES
☐ COLOR OR BLACK AND WHITE PHOTOGRAPHS
☐ GRAY SCALE DOCUMENTS
☐ LINES OR MARKS ON ORIGINAL DOCUMENT
☐ REFERENCE(S) OR EXHIBIT(S) SUBMITTED ARE POOR QUALITY
OTHER.

### IMAGES ARE BEST AVAILABLE COPY.

As rescanning these documents will not correct the image problems checked, please do not report these problems to the IFW Image Problem Mailbox.